

AgRISTARS

TM83918

RR-G2-04258

APRIL 1982

A Joint Program for
Agriculture and
Resources Inventory
Surveys Through
Aerospace
Remote Sensing

Renewable Resources Inventory

Final Report

Detecting Forest Canopy Change Using Landsat

(NASA-TM-83918) DETECTING FOREST CANOPY
CHANGE USING LANDSAT Final Report (NASA)
83 p

N89-70808

by

Ross F. NEISON

Unclas
00/43 0189943

NASA/Goddard Space Flight Center
Earth Survey Applications Division
Earth Resources Branch
Greenbelt, MD 20771



NASA



TM 83918

DETECTING FOREST CANOPY CHANGE USING LANDSAT

Ross Nelson

April 1982

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland 20771

Abstract

Multitemporal Landsat multispectral scanner data were analyzed to test various computer-aided analysis techniques for detecting significant forest canopy alteration. Three data transformations, differencing, ratioing, and a difference of ratios, were tested to determine which best delineated gypsy moth defoliation. Response surface analyses were conducted to determine optimal threshold levels for the individual transformed bands and band combinations. Results indicate that, of the three transformations investigated, a difference of ratios ($\frac{\text{Band7}}{\text{Band5}}$) transformation most accurately delineates forest change due to gypsy moth activity. Band 5 (0.6-0.7 μm) ratioed data did nearly as well, however, other single bands and band combinations did not improve upon the band 5 ratio results.

Detecting Forest Canopy Changes Using Landsat

Abstract	iii
I. Introduction	1
A. Change Detection in Forestry	1
B. The Gypsy Moth and Forest Canopy Change	2
II. Literature Review	5
A. Delta Classification	5
B. Analysis of Multitemporal Data Simultaneously	9
1. The Use of Untransformed Landsat Data	11
2. The Use of Transformed Landsat Data	11
C. Techniques Selected for Study	16
III. Materials	19
IV. Procedure	25
A. Data Transformations	25
1. Differencing	25
2. Ratioing	25
3. Difference of Ratios	26
B. Data Analysis, Techniques used to Alarm Changed Pixels	26
C. Accuracy Considerations	28
D. Detecting Change using Individual Bands	32
E. Modeling Combined Accuracy - Single Band Images	33
F. Multiband Analysis--Difference and Ratio Images	34
V. Results	37
A. Analysis of Individual Bands	37
1. Difference Image	38
a. Empirical Results - Difference Image	38
b. Modeling - Difference Image	41
2. Ratio Image	44
a. Empirical Results - Ratio Image	44
b. Modeling - Ratio Image	48
3. Difference of Ratios	48
a. Empirical Results - Difference of Ratios Image	48
b. Modeling - Difference of Ratios Image	49
B. Multiband Image Analysis	51
C. Comparison of All Approaches	53
D. Characteristics of Landsat Data and Transformations	54
VI. Conclusion and Discussion	63
VII. Appendix A and B	73

Detecting Forest Canopy Change Using Landsat

Ross Nelson

I. INTRODUCTION

A. Change Detection in Forestry

Foresters have traditionally relied on airphotos, photo interpreters, and ground sampling to keep their forest inventory records current. The inventory techniques, though effective, are cumbersome, and catastrophic events can quickly change the quantity and condition of the wood resources available. Fire, wind throw, ice, and insect damage may profoundly affect a manager's decisions concerning harvesting, silvicultural, and pest control practices in his district. Change detection techniques which quickly and accurately delineate forest canopy alteration provide the information necessary to make intelligent management decisions.

Satellite data may be used to detect gross forest canopy alterations. Numerous studies have shown that broad cover types--conifer, hardwood, agricultural areas, water, etc.--may be consistently and reliably classified using satellite data. Satellite remote sensing, however, does have limitations, and various studies have demonstrated an inability to consistently and accurately identify tree species, crown density differences, and age class differences. These characteristics may be critically important in terms of an assessment of the impact of canopy change in a forest.

Due to the limitations of the Landsat data, some have suggested that Landsat may best be utilized as the first stage in a multistage sampling design (Heller, 1978; Smith, 1979). The broad cover types should be reliably delineated using satellite data. In terms of forest canopy alteration, change and unaltered areas should be delineated. Finer resolution systems and ground observations can provide the details necessary for the management decisions.

This investigation assesses the capabilities of the Landsat data and computer-aided analysis techniques to provide such first stage change information. An area attacked by the gypsy moth was studied to determine how well selected change detection procedures delineated defoliated forest from unaffected forest.

B. The Gypsy Moth and Forest Canopy Change

The gypsy moth was introduced into the United States in 1869 when a French scientist imported the egg clusters for breeding experiments with the silkworm. In the course of attempts to produce a hearty, commercial hybrid, gypsy moth eggs or larvae were lost; the import from France quickly established itself. By 1890, infestations encompassed a 350 mile area around Medford Massachusetts. Since that time, in spite of chemical and biological controls, the gypsy moth has spread to northern Maryland, southern Maine, and from the Atlantic Ocean to the Ohio-Pennsylvania border (see Figure 1). It defoliates many of the northern hardwoods, favoring the oaks, birches and aspen; conifers may also be attacked if the epidemic is severe. The moth kills or weakens timber on hundreds of thousands of acres a year. Control practices, which include the application of pesticides, pheromones, microbial or viral insecticides, and the release of natural predators or sterilized male moths, depend on the accurate location of areas supporting the epidemic populations. These areas are currently located by ground egg mass counts or by aerial sketchmapping. During the past decade, the utility of Landsat satellite data for monitoring the distribution and severity of defoliation has been investigated. Research has shown that the digital analysis of Landsat data cannot delineate all defoliation types of interest. Though a heavily defoliated canopy can easily be distinguished from all other forest cover types, the spectral and spatial resolutions of the Landsat satellites are such that a healthy forest canopy cannot be accurately distinguished from forests in which 30-60% of the

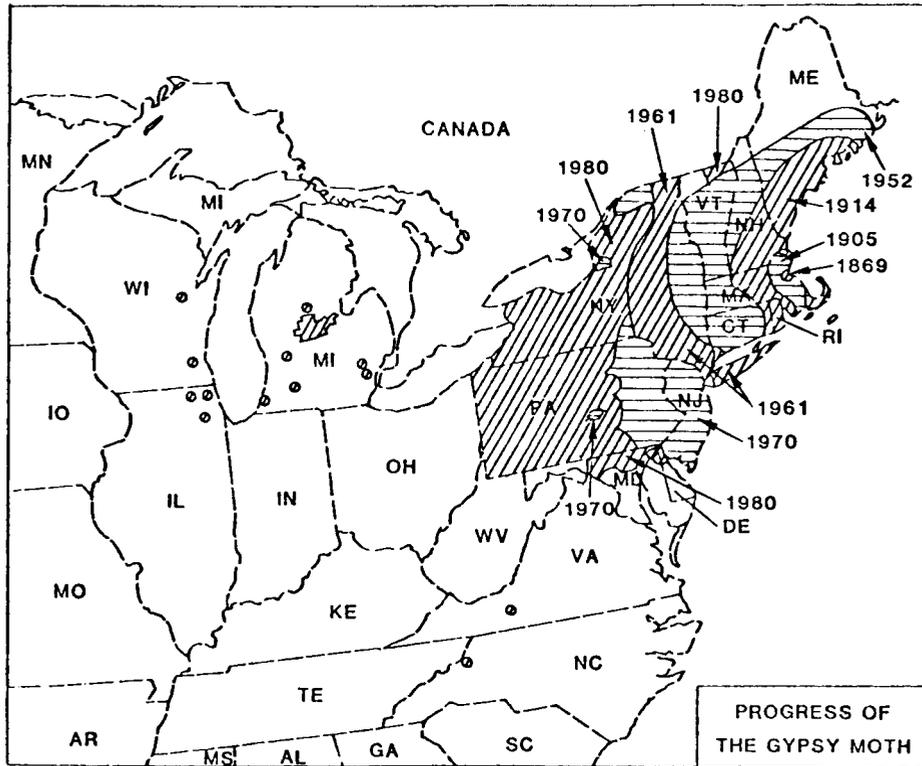


Figure 1 Expansion of the gypsy moth range since its introduction in 1869 (from Marshall, 1981).

canopy has been removed by the gypsy moth (Williams and Stauffer, 1979; Nelson, 1981; Williams and Ingram, 1981). These moderately defoliated areas support potentially explosive gypsy moth populations; as such these are prime targets for control measures.

The usefulness of the Landsat satellite data then, lies in the capability to assess relatively severe damage, that is, to detect significant forest changes. The objective of this research was to determine which of three change detection approaches most accurately portrayed forest canopy alterations due to gypsy moth defoliation. The three data transformations, image differencing, image ratioing, and the difference of ratios, were investigated to determine those band combinations and threshold levels which maximized change classification accuracy.

II. LITERATURE REVIEW

Digital change detection approaches may be characterized by (1) the data transformation procedure (if any), and (2) analysis techniques used to delineate areas of significant alterations. A tabular breakdown of a variety of change detection approaches is given in Table 1, along with references to those who have used that particular approach. The purpose of the table is twofold. First, it is a concise summary of a majority of the change detection work done to date. Second, it gives the reader an idea of the number of approaches available for digitally detecting land cover alterations.

Figure 2 outlines a series of steps which an analyst may take in order to implement a particular change detection approach. Many of the authors listed in Table 1 are included in Figure 2 so that the reader might better understand the particular approach taken by these investigators. The steps depicted in this Figure are separated into two distinct sections by the first decision node. At this point the analyst must decide upon one of two basic methods for detecting change; 1. delta classification or 2. the analysis of multitemporal data as a single data set. Delta classification involves the classification of the individual dates of a multitemporal data set into land cover types and a subsequent comparison of the land cover classifications. Changes in the identity of a particular pixel from time one (t_1) to time two (t_2) connote land cover alteration. The second basic approach calls for the simultaneous analysis of multitemporal data. Using this approach, the multitemporal data is classified to identify changed areas. Each of these procedures is discussed more fully in subsequent sections.

A. Delta Classification

A standard change detection procedure investigated by many has been the delta classification procedure. Delta classification change detection relies

Table 1. Change detection research broken down by (1) the data transformation used (if any) and (2) the analysis technique used to detect change.

Analysis Technique Used to Detect Change	Two Dates Analyzed Simultaneously							Delta-Classification 2 Dates Analyzed Ind.	
	Raw Data	Transformed Data					Correlation		Raw Data
		Difference	Ratio	Vegetative Index Difference	Regression	Greenness-Brightness			
S.D Threshold (Density Slice)		Stauffer and McKinney (78) Ingram et. al (81) Toll et. al (80) *	Todd (77)	Angelici et. al (77)	Ingram et. (81)	Byrne and Crapper (80) Toll et. al (80)	Coiner (80)		
Supervised	Williams and Haver (76)	Anuta and Bauer (73)						Gordon (80)	
Modified Supervised		Anuta (74)						Colwell et. al (80)	
Unsupervised	Spectral	Barthmaier et. al (80) Weismiller et. al (77)						Swain (76) Joyce et. al (80) Weismiller et. al (77) Riordon (80)	
	Spectral-Spatial					Colwell et. al (80)			
Layered	Weismiller et. al (77)								
Visual Interpretation	Colwell et. al (80)								

*'s mark the techniques tested in this study.

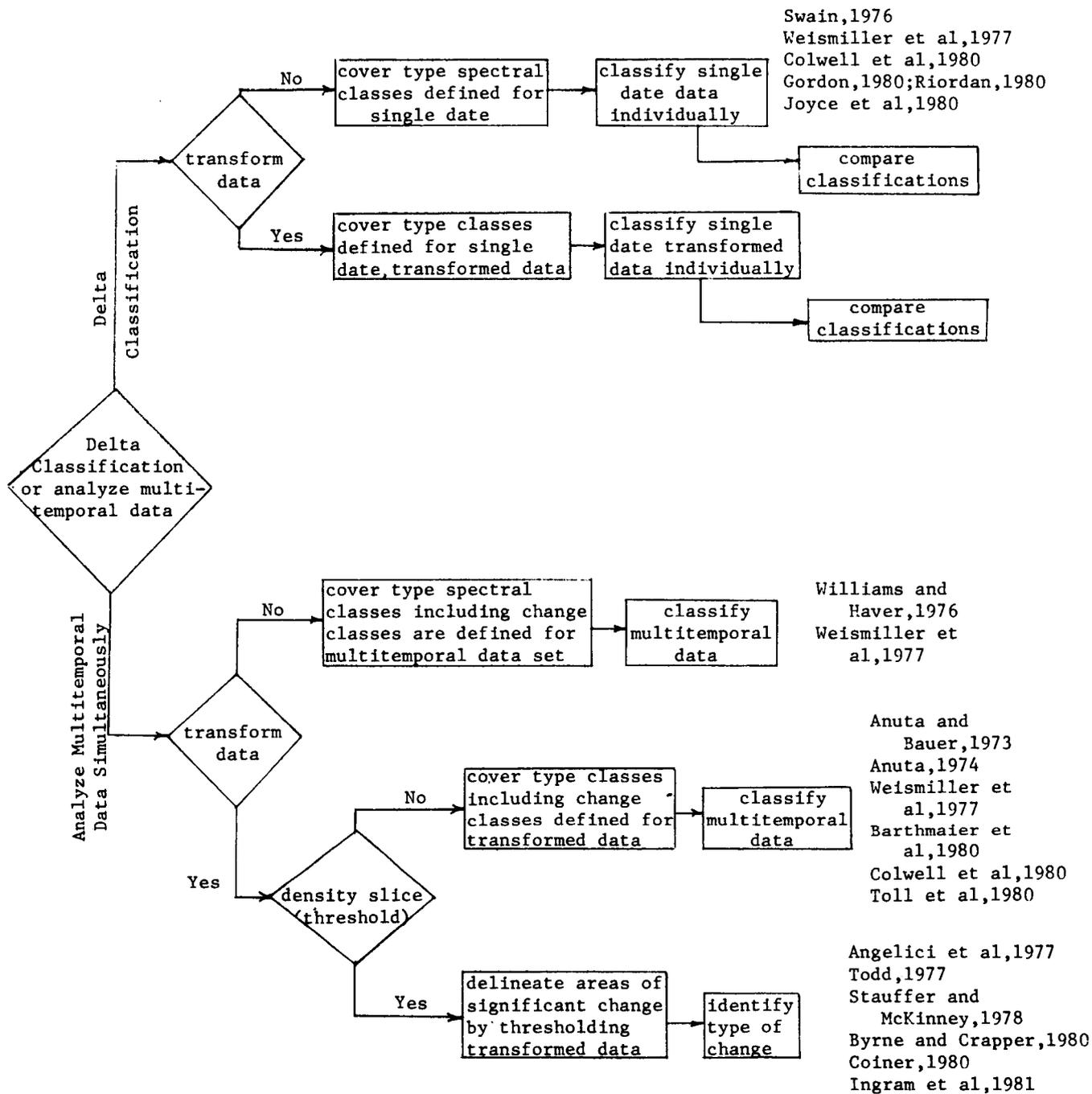


Figure 2 Approaches for detecting change.

on the independent, correct classification of two data sets. Inaccuracies in either lead to spurious change when the classifications are compared. Swain (1976) endorsed a delta classification approach, however he employed a processor which reduced the spurious change problem. Swain compared the capabilities of a per-point and a per-field classifier (ECHO, Kettig and Landgrebe, 1975) and found that the ECHO classifier reduced speckling, i.e., the random incorrect classifications of individual pixels. He concluded that:

"urban encroachment of agricultural areas can be detected through classification of multitemporal (Landsat) data. However, the size of the change areas must be large relative to the resolution of the sensor, and the classifications must be done with relatively high accuracy with respect to the classes that will show the change."

Weismiller et al., (1977) also claimed that the delta classification technique reliably identified areas of change. This conclusion is based on a qualitative comparison of the change results; no coincident ground information was available.

Riordan (1980) produced unsupervised classifications of 1973 and 1978 Landsat MSS data and compared the classifications to detect nonurban to urban change. She reported a 67% accuracy figure. Gordon (1980) used Landsat data and a delta classification approach to monitor land-use change in Ohio. Based on a rigorous quantitative assessment, he wrote:

". . . we must conclude that substantial errors are associated with the use of Landsat data for land cover and change analysis."

The poor showing of the delta classification approach may be attributed to the individual data sets' poor classification accuracies and possibly, to mis-registration errors.

The delta classification approach was not considered as a viable alternative in this study because: 1) it depends on two fairly accurate, independent land cover classifications and 2) the two classifications must be comparable. Toll et al., (1980) noted that the poor performance of the delta classification approach may, in part, have been attributable to "the difficulty of producing comparable

classifications from one date to another". As noted previously, an error in either classification results in an error in the change image. Spurious, small areas of change result, which may cause the amount of change to be overestimated.

B. Analysis of Multitemporal Data Simultaneously

The disadvantages of the delta classification procedure may be avoided by analyzing the multitemporal data simultaneously. When a multitemporal data set is analyzed as a single multiband image, the analysis sequence may be characterized by the method used to identify change. Essentially two change detection classification procedures are available, multitemporal classification and density slicing, also called thresholding.

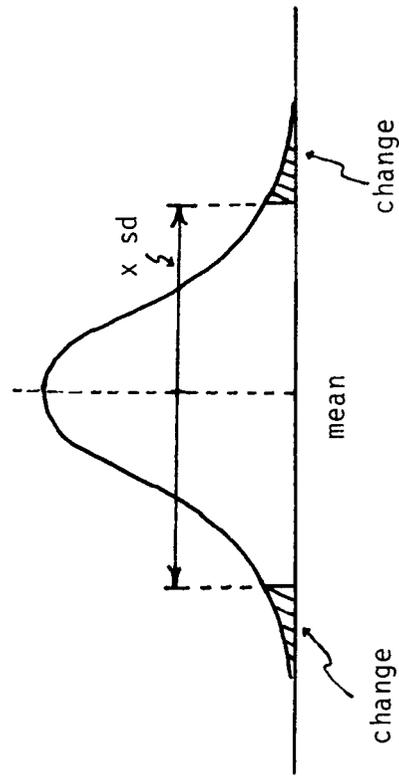
Multitemporal classification techniques may be applied to raw or transformed Landsat data. The analyst develops training statistics which mathematically describe the cover types of interest, including change categories. These statistics are used to classify the study area into changed and unchanged land cover types.

Density slicing or thresholding techniques may only be applied to transformed data. The transformations are necessary to provide a framework wherein large departures from the norm may be recognized. Transformations are done on a pixel by pixel basis within a given band. The spectral response of the same piece of ground is compared at two different times to see if any gross spectral changes have occurred in that time interval. Hence relatively large deviations from the mean indicate land cover alteration. Changed pixels may be thresholded to one side of the mean or to both sides, depending on the spectral characteristics of the change of interest and the data transformation used (see Fig. 3).

The transformations which are commonly used in change detection work and the research results of investigators using those transforms are discussed below. The analysis technique used to highlight change is also given within

CHANGE DETECTION: THRESHOLDING

BOTH SIDES OF MEAN



ONE SIDE OF MEAN

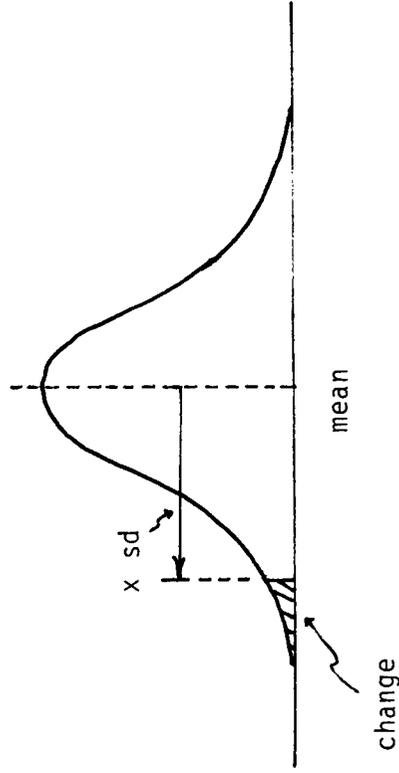


Figure 3 Thresholding transformed data to detect change. The mean and standard deviation of a given band of transformed data are calculated. Those pixels whose transformed values lie beyond a certain threshold (in this case, x standard deviations, or sd) are considered changed.

the context of the characteristics of the data. Again, in general, the analyst has only two choices--multitemporal classification or thresholding, and thresholding can only be used if the multitemporal data has been appropriately transformed.

1. The Use of Untransformed Landsat Data

Williams and Haver (1976) and Weismiller et al. (1977) developed training statistics for various change classes using untransformed Landsat data. Williams and Haver (1976) identified drastic forest canopy alterations (e.g., fire damage and clearcutting) and suggested that the use of multitemporal data allows rapid updates of harvesting activities or catastrophic events. Weismiller et al. (1977) concluded that the multitemporal classification approach has "undeveloped potential". In his particular investigation, however, this approach did not agree with the results of a delta classification which he used as a reference source. Though the accuracy assessment was qualitative, Weismiller found that seemingly stable areas were often noted as changed using the multitemporal classification approach. This work suggests that the multitemporal classification approach using raw Landsat data might best be limited to delineating changes which are spectrally quite distinct.

2. The Use of Transformed Landsat Data

Landsat data may be transformed for any one of three reasons: (1) to accentuate various land cover features or change phenomena; (2) to reduce data dimensionality; and/or (3) to enable use of the thresholding technique to delineate change. The transformations used by a variety of researchers to detect change are reviewed below.

Differencing: A differenced data set is formed by subtracting the greylevel value of a particular pixel in a given band at time one from the corresponding greylevel at time two. Large positive or negative values connote change (see Fig. 3). The transformation also serves to reduce an eight channel data set to four channels.

A number of researchers have used the differencing technique to detect change, primarily in urban environments, with varying degrees of success. Toll et al. (1980) investigated image differencing, principal components transformation, and delta classification to see which most accurately portrayed conditions in Richmond, Virginia and in Denver, Colorado. Training statistics were developed for the differenced and principal components data, and the change categories were classified. The differenced data produced the highest overall change detection accuracies. Anuta (1974) classified differenced data using a supervised approach in an attempt to identify new construction sites. He found that construction could be accurately identified, but many other types of change, such as agricultural and industrial, were incorrectly included in the construction class. Anuta's major concern was not that areas that had changed were falsely alarmed, but that the identification of the type of change was incorrect. Ingram et al. (1981) found that simple image differencing and thresholding produced urban change results (Denver) as good as those obtained using much more sophisticated approaches. Finally, Barthmaier et al. (1980) reports a quasi-operational use of differenced Landsat data and supervised classification techniques to detect clearcutting operations in the state of Washington. The areal tallies compiled from the Landsat data for various state or private tracts are compared with the tallies turned into the Department of Revenue for tax purposes. Obvious inequities are investigated by the Department of Revenue.

Ratios: Ratioing is similar to image differencing in that both approaches compare relative reflectance measurements at two different times. However, greylevels are divided rather than subtracted to produce the transformed 4 band image.

Data ratioing has not been as intensively investigated as image differencing, perhaps because of software and hardware limitations (ratioing

results in a real numbered image, not a byte image). Todd (1977) ratioed band 5 pixel values from two different years to determine urban change in Atlanta, Georgia. Only ratios to the low side of the mean were considered changed (threshold approach). He then classified the most recent MSS data set to determine the type of change that had occurred. His overall evaluation indicated that 91.4% of all land use and land cover change was correctly identified. This included 78% of the total number of change areas. Accuracies decreased when attempts were made to identify the types of change. Todd noted that omission and commission errors occurred, for the most part, in relatively small areas.

Difference of Ratios: The difference of ratios (in this study, (Band 7/Band5, t_1) - (Band 7/Band 5, t_2) for a given pixel) is fundamentally different from the first two transformations in that vegetation density measures are compared. Landsat bands 5 and 7 are well situated for monitoring green vegetation; the 0.50-0.60 μm region (band 5) is centered on the red absorption wavelengths of a green canopy, and the 0.80-1.10 μm reflective infrared region is highly reflected by vegetation due to the internal structure of the leaf (Tucker and Maxwell, 1976). The ratios of the reflectance measurements in the red wavelengths to those in the photographic IR have been used by researchers as a green biomass indicator. Radiometer studies (Jordan, 1969; Pearson and Miller, 1972; and Tucker, 1979), aircraft scanner studies (Stoner et al., 1972), and Landsat MSS data studies (Maxwell, 1976; Justice, 1978; and Nelson 1981) have shown that the IR/red response ratio is sensitive to the amount of green leaf biomass being sensed. Comparing the 7/5 ratio between dates would provide an avenue for deciding whether or not a vegetation canopy has been significantly altered. Only one study (Angelici et al., 1977) was found which used the difference of ratios data and the thresholding technique to delineate

changed areas. Unfortunately, as seems to be the case with a majority of change detection studies, no quantitative assessment of the results was performed. Hence, the capabilities of such a technique are as yet undocumented.

Regression: Regression involves an approach where a particular data value at t_1 is used to predict the corresponding data value at t_2 using a simple linear relationship. Assuming that most of the image has not changed, those pixels which have large residuals have most likely undergone change. Ingram et al. (1981) compared the change detection capabilities of the regression data to a number of other data transformations. Thresholding was used to detect change in the various data sets. They found that, while the regression transform did do better than more sophisticated transforms, it did not do as well as simple image differencing.

Greenness-Brightness: Colwell et al. (1980) used a vegetation-soil index to monitor forest canopy changes in South Carolina. The data transformation, called TASCAP, is the first step in a complete change detection system (TASCAP/BLOB/CVA) developed by Colwell et al. and reviewed below.

TASCAP refers to the initial data transformation which converts an eight channel, multitemporal data set into a four channel greenness-brightness data set. Each date is characterized by two channels, a greenness channel related to the amount of vegetation in the field of view, and a brightness channel related to the contribution of soil (Kauth and Thomas, 1976).

This four channel, transformed data set is clustered using a spectral-spatial clustering algorithm called BLOB (Kauth et al., 1977). Spectrally similar, contiguous areas are found in the data. Each blob (spectral-spatial cluster) vector has four components, consisting of the means of the greenness and brightness values for the two dates. The greenness and brightness measures should be similar in the areas (blobs) that do not represent change. Blobs

formed over change areas should vary significantly in the transformed channel values.

The magnitude of the differences between the greenness and brightness measures for each date determines whether or not the change is significant, i.e., worth considering. The directions of the changes in these values determine the type of land cover alteration that has taken place. Blobs whose greenness values decrease and brightness values increase have lost vegetation, while the converse may indicate reforestation or field crop development. The thresholds and directions have been defined in an approach called Change Vector Analysis (CVA) (Malila, 1980).

Colwell et al. (1980) used the TASCAP/BLOB/CVA technique to detect changes in Kershaw County, South Carolina. No ground reference accuracy assessment was done, but the results of the TASCAP/BLOB/CVA approach and a delta classification approach were compared to a dot grid estimate of change derived from a Landsat multitemporal color composite. Colwell found that the delta classification approach resulted in identifying six times more change than was indicated by the dot grid estimate, while the TASCAP/BLOB/CVA results and dot grid estimates were comparable. They wrote that the delta classification overestimate of the change area was most probably due to:

"(1) spurious recognition of anomalous single pixels in delta classification, which is reduced in BLOB/CVA, and (2) inconsistent date to date classification of the same pixel in delta classification, which is eliminated by joint analysis of both dates of data in BLOB/CVA." (Colwell et al., 1980, pg. 51)

Principal Components: This transformation defines the major axes of variation within the data set and mathematically rectifies the data to lay along the axes. The variance components are orthogonal and may hold interesting discriminatory information. Toll et al. (1980) investigated three change detection procedures, image differencing, principal components transformation

prior to differencing, and a delta classification approach. They used thresholding techniques to define change areas and found that image differencing produced the highest accuracies.

Disagreement exists over which principal component is most useful for discriminating various land cover types. All agree that the first component contains variability due to the overall scene brightness. Robinson (1979), Riordan (1980), and Friedman (1979) have suggested that the second or third components contain the variability useful for cover type discrimination. Toll et al. (1980) refutes the usefulness of the 2nd component of a single date data set, at least in the context of urban-nonurban discrimination. Byrne and Crapper (1980) report that the third and later components of a multitemporal data set hold the most interesting information for change detection purposes. The first two components reflect the variability due to scene brightness and the presence of clouds. Obviously, care must be exercised when selecting the proper principal component to discriminate various land cover features.

Other data transformations have been developed for specific land cover purposes (predominantly for use in agricultural situations). Many of these are described in Williams and Stauffer (1979) or Tucker (1979). The single guiding factor in selecting an appropriate transformation, if one is necessary at all, is that the transformed data should accentuate the change phenomenon of interest.

C. Techniques Selected for Study

Three data transformations, differencing, ratioing, and a difference of ratios were selected for investigation. The difference transformation is selected for two reasons. First, it was the most widely used transformation, making it a useful metric against which the results of other transformations

may be compared. Second, Toll et al. (1980) and Ingram et al. (1981) found that differenced data produced results as good or better than more sophisticated transformations. The ratio transformation was investigated to determine the impact of a mathematical operand on the same data sets. Differences in the change detection capabilities of difference or ratioed data would be due solely to the fact that one transformed data set was produced by subtraction, the other by division. Hence comparison of the results of differenced and ratioed data would quantify the effects of different arithmetic operations. The difference of ratio transformation was studied to see if change in a vegetative index (i.e., a value responsive to the amount of green vegetation in the field of view) characterized forest canopy change better than relative reflectance comparisons (i.e., the first two data transformations).

Thresholding techniques are used to delineate change. The thresholding or density slicing approach is one that requires no a priori information once the threshold is set. Multitemporal classification or delta classification, on the other hand, always requires information so that training statistics can be developed for the classifiers. Hence, from an operational stand point, the thresholding technique is potentially much more useful. If one can define the optimal data transformation, wavelength(s), and the thresholds, and if these thresholds are relatively stable from year to year, then an operational change detection system may be realized.

The three data transformations were tested to see which was most useful for delineating forest canopy alteration. Threshold limits were defined to maximize change detection accuracy. Throughout the study, the Landsat data were processed to define only two forest cover conditions, change or no change.

III. MATERIALS

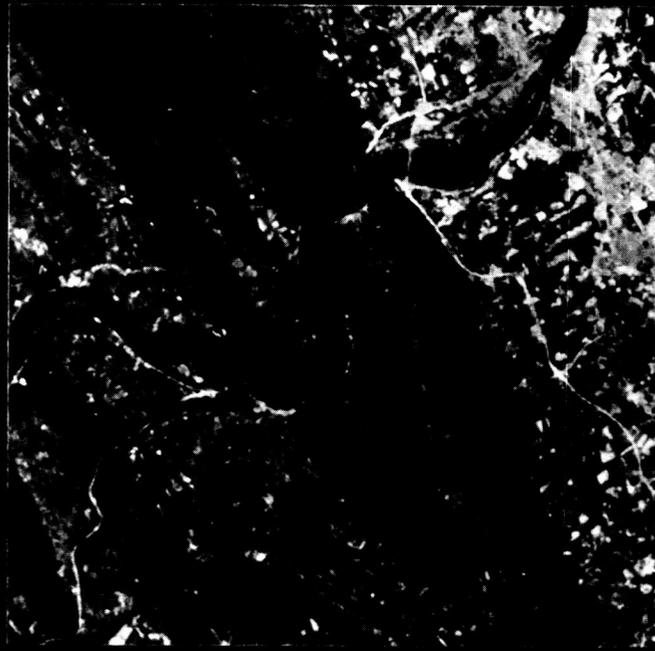
A. Landsat Data

July 19, 1976 (2444-15001) and June 27, 1977 (2887-14520) Landsat scenes obtained over Harrisburg, Pennsylvania (path 17, row 32), were geometrically corrected, registered and resampled to a 50 meter grid (see Fig. 4). A 286 line by 217 sample subsection corresponding to the Wertzville, Pennsylvania 7-1/2 minute USGS quadrangle map was chosen for analysis. The quadrangle, encompasses mountainous areas which are thickly forested, predominantly an oak-hickory cover type. These forests were not defoliated in 1976, but were extensively damaged by the gypsy moth in 1977. The June 27, 1977 data set was obtained during the peak defoliation period, which occurs from mid June through early July.

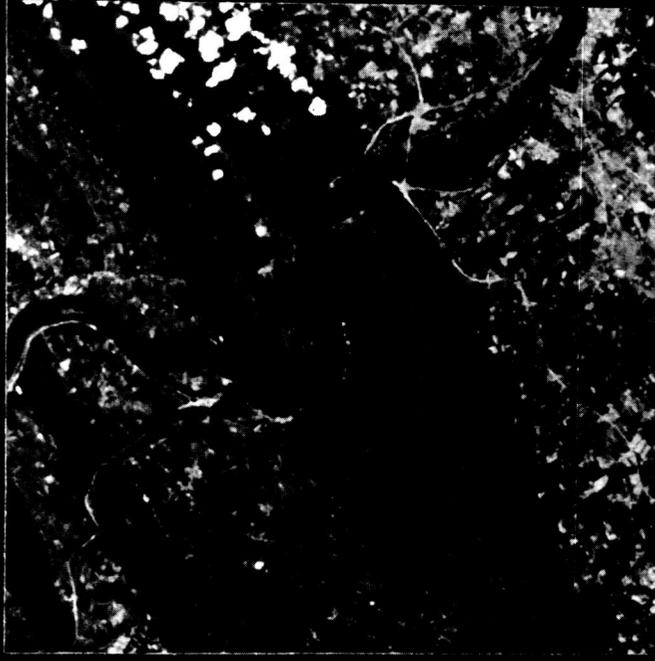
Color infrared aerial photography obtained on June 24, 1977 was available over the entire quadrangle. The 1:48,000 scale airphotos (see Fig. 5) were used to delineate areas of heavy (60-100% leaf removal) and moderate (30-60% leaf removal) defoliation. Acetate was laid atop the photos and the defoliated areas were outlined. This information was transferred to the 7-1/2 minute quad map (Fig. 6) using a Zoom Transfer Scope. The defoliation boundaries were then digitized from the quad map using the HP-3000 Geographic Entry System (Stauss et al., 1978).

The digitized defoliation information was combined with a Landsat generated forest/non-forest mask to form the ground reference image (GRI) to which all change detection products were compared. The forest/non-forest mask was generated from the July 19, 1976 (i.e., healthy) Landsat data set. Supervised forested training statistics, only, were input to a Bayesian classifier. The classifier generated a confidence map that assigned a probability to each pixel that the pixel belonged to the forest class. This confidence map was density sliced to separate the forested and

**LANDSAT SUB-IMAGES OF THE HARRISBURG, PA. AREA
SHOWING AN INCREASE IN GYPSY MOTH DEFOLIATION BETWEEN
1976 AND 1977**



JULY 19, 1976



JUNE 27, 1977

10km/6.25 Mi.

NOTE: THESE IMAGES HAVE BEEN CONTRAST ENHANCED, GEOMETRICALLY CORRECTED AND REGISTERED TO ONE ANOTHER

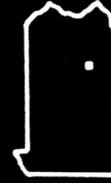


Figure 4 The Landsat data used in the change detection analysis.

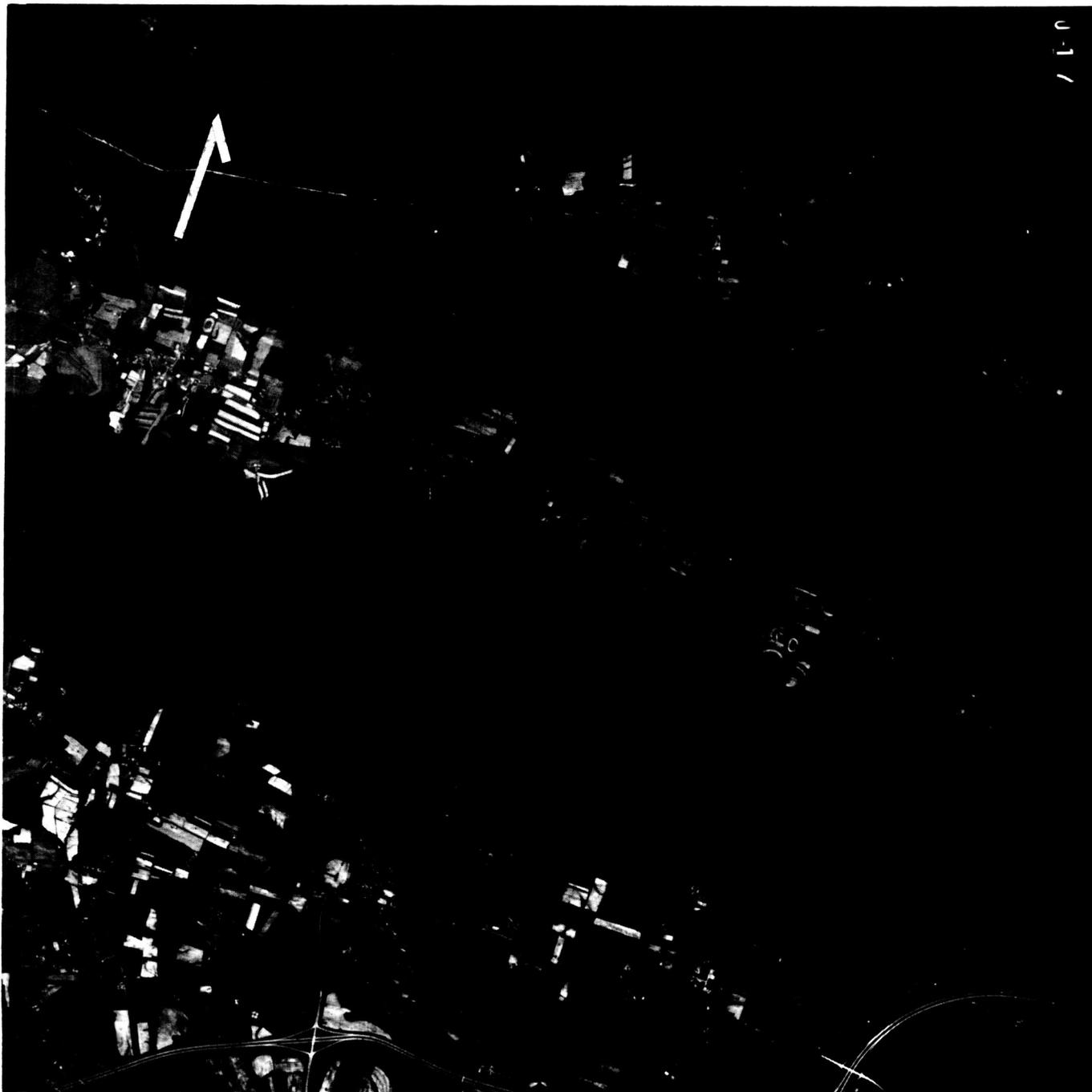


Figure 5 One of the 1:48,000 scale color infrared airphotos used to delineate defoliation conditions on the Wertzville Quadrangle, Pennsylvania. The registration marks (black cross hairs) permit scale comparison with the ground reference image (Figure 7, lower left corner).

non-forested areas. A stratified random sample of 230 pixels (115/strata) indicated that the mask's accuracy was 89.95% +5% at the 95% level of confidence.

Combining the forest/non-forest mask with the digitized defoliation information resulted in a four class ground reference image (GRI): 0 - non-forest, 1 - heavy defoliation, 2 - moderate defoliation, 3 - healthy forest (see Figure 7). Any discrepancies between the defoliation data and the forest/non-forest mask were rectified in favor of the mask. In other words, if the defoliation information showed that a given pixel was moderately defoliated, but the mask showed it as non-forest, then the mask was assumed correct. That pixel would be non-forest in the GRI. The Interactive Digital Image Manipulation System (IDIMS) software (Electromagnetic Systems Laboratory, 1978) was used to manufacture the GRI.

IDIMS was also used to generate the transformed data sets from the raw multitemporal Landsat data. In addition, the image processing system was used to generate change images from these transformed data sets using thresholding techniques. The specific processes involved are explained in the PROCEDURES section.

The change products resulting from the thresholding of the transformed data were compared to the GRI to assess the accuracy of the change classification. The accuracy evaluations were done using a software package called ASSESS2. ASSESS2 compares two byte images and quantitatively evaluates their similarity. One of the two images is designated the ground reference image (in this case the GRI); the other, the classified image, is compared to the former. The software evaluates the accuracy of classification on a per pixel basis and by analyzing polygons (Chaiken, 1979). Only per-pixel accuracies were considered in this study.

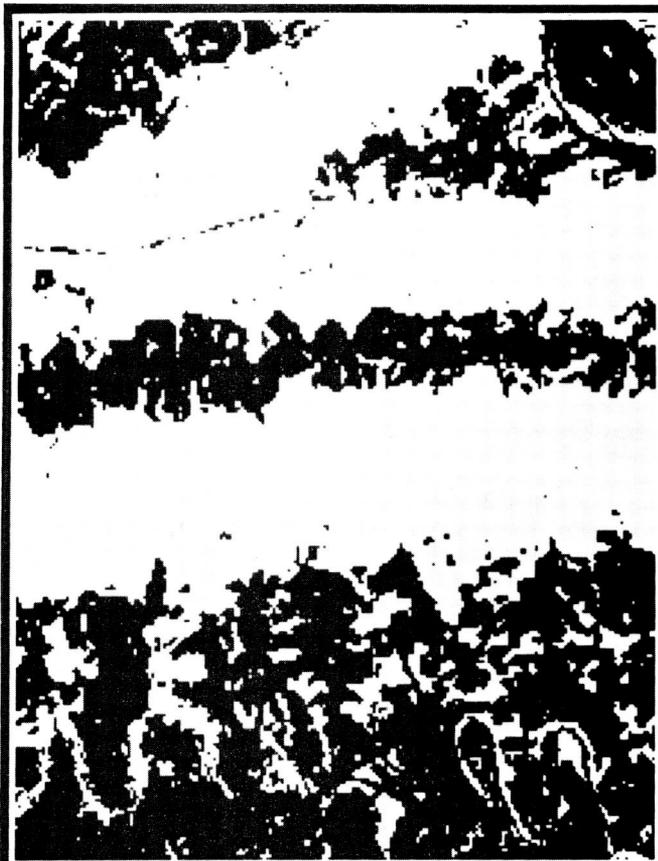


Figure 7 The ground reference image was manufactured by combining the forest/nonforest mask and the digitized defoliation information.

- Upper left: forest/nonforest mask
 - white - forest
 - black - nonforest

- Upper right: digitized defoliation information
 - light grey - moderate defoliation
 - dark grey - heavy defoliation

- Lower left: ground reference image (GRI), colors as noted above

IV. PROCEDURE

Any change detection approach can be categorized by (1) the data transformation method; and (2) the analysis technique used to detect change areas. The transformations and techniques used in this study are described below.

A. Data Transformations

Three data transformations were investigated to determine which most accurately detected gypsy moth defoliation.

1. Differencing: An eight band multitemporal Landsat image was reduced to a four band differenced image. A given differenced band is formed by subtracting the pixel value at time two (t_2) from the corresponding pixel value at time one (t_1):

$$D_{ijk} = X_{ijk,1976} - X_{ijk,1977} + 127$$

where: i = line coordinate
 j = sample coordinate
 k = band number (MSS 4-7)
 X = greylevel value for that pixel
 D = differenced value
127 = constant added to produce a non-negative image assuming a dynamic range of 128 greylevels.

The differenced image values, then, may range from 0-254, with an expected mean of 127 if there is no change between images.

2. Ratioing: Again, an eight band image was reduced to a four band ratioed image, the ratioed bands are manufactured thusly:

$$R_{ijk} = \frac{X_{ijk,1977}}{X_{ijk,1976}} \text{ where all terms are as previously defined.}$$

R = ratioed value of the given pixel.

Note that the differenced image is a byte, or discrete, image. The ratioed image is a real image with an expected mean of 1.00 for a given band if no change has taken place. The potential range of this image is 0.00 to positive infinity.

3. Difference of Ratios: An eight band image was reduced to a one band data set using the following transformations.

$$DR_{ij} = \frac{X_{ij7,1976}}{X_{ij5,1976}} - \frac{X_{ij7,1977}}{S_{ij5,1977}} + C$$

(healthy) (defoliated)

where: i,j: have been previously defined
 7,5: refers to the band numbers, 5 (0.60-0.70 μm), 7 (0.8-1.1 μm).
 C : is an arbitrary constant added to produce a non-negative real image.

In the case of the data used for this study, C=4.0. The potential range of this image is 0.00 to positive infinity. A simpler method of depicting this transformation follows:

$$\text{Difference of Ratios} = \frac{\text{Band7}}{\text{Band5}}_{\text{time 1}} - \frac{\text{Band7}}{\text{Band5}}_{\text{time 2}} + 4.0$$

Two factors should be noted concerning the characteristics of these data transformations. First, the differenced image is a four band byte, or discrete, image. Its' dynamic range is the integer values between 0 and 254 inclusive. The Ratio and Difference of Ratios images are real, continuous images. Second, the Difference and Ratio transformations result in images which are basically albedo comparisons at time 1 and at time 2. The Difference of Ratios image is a comparison of green biomass measures at time 1 and 2.

B. Data Analysis, Technique Used to Alarm Change Pixels

The data were transformed and the mean and standard deviation of each transformed band were calculated. Various standard deviation threshold levels were tested for any particular transformed band to see which threshold produced the highest change classification accuracy. An example may serve to explain how the thresholding was done.

Example: Develop a change classification image for differenced band 5 data. Threshold: 1.5 standard deviations.

The mean and standard deviation of the difference band 5 image may be calculated. Mean: 125.765 SD: 3.718

Calculate the threshold values for 1.5 standard deviation: $125.765 + (1.5 \text{ stan. dev}) (3.718) = 126.765 + 5.577 = 132.342$. Hence anything further than 1.5 standard deviations from the mean is considered change. However, due to the expected spectral change of the defoliated forest canopy in this wavelength region, only difference values to the low side of the mean (in this case, below 121.188) were considered change due to gypsy moth activity. So, band 5 difference values less than or equal to 121 were flagged as change.

Pixels regarded as unchanged at this particular standard deviation level (i.e., difference band 5 values of 122 or greater) were set equal to 0. Changed pixels (values equal to or below 121) were set equal to 1. This 0/1 change classification was compared to the GRI to see how well it correctly identified the gypsy moth damage.

The use of the standard deviation criterion to delineate changed pixels in no way assumes a normal data distribution. In fact, it was assumed that the data would be markedly non-normal. Lack of data normality merely prevents the experimenter from using the Empirical Rule which states that 68% of the population lies within 1 standard deviation of the mean, 95% within two standard deviations, etc. The standard deviation threshold criterion is merely one method of determining the actual cutoff value above or below which pixels are noted as changed.

A standard operating procedure in most studies which have utilized the thresholding approach has been to calculate threshold values on both sides of the mean, since all types of change were of interest. Generally threshold

levels on the order of 2 to 3 standard deviations were evaluated since the experimenters were looking for a relatively small amount of total scene change. This study was interested solely in change in the forest canopy, more specifically, change due to gypsy moth defoliation. If a forest canopy is defoliated, the change in the spectral response of that canopy over time is predicatable. Nelson (1981) showed that a canopy which has suffered gypsy moth defoliation generally increases its reflectance in the two visible Landsat bands (4 and 5) and exhibits a decrease in the two infrared Landsat bands (6 and 7). This information was used to threshold the transformation images to one side of the mean only, as follows:

Difference Image	Band 4 (0.5 - 0.6 μm)	low side of mean
	Band 5 (0.6 - 0.7 μm)	
	Band 6 (0.7 - 0.8 μm)	
	Band 7 (0.8 - 1.1 μm)	
Ratio Image	Band 4	high side of mean
	Band 5	
	Band 6	low side of mean
	Band 7	
Difference of Ratio Images	Since one would expect the 7/5 ratio to decrease with increasing defoliation, and the 1977 ratio (defoliated) was subtracted from the 1976 ratio; this single band image was thresholded to the high side of the mean.	

Hence, pixels on the tail end of the distribution not considered were not noted as changed, no matter how "abnormal" the data value. Though there may have been some legitimately changed pixels in these tails, the change most likely was not due to gypsy moth defoliation. Only canopy alterations due to the gypsy moth infestation were of interest.

C. Accuracy Considerations

Two major questions concerning classification accuracy had to be addressed prior to or during the study. The first dealt with those cover types which would be considered "change" in the ground reference image; the second concerned a suitable accuracy criterion which could be maximized. The problem

and the solution to each problem are given below.

The GRI contained 3 forest classes, healthy, moderately defoliated, and heavily defoliated forest. The 1976 Landsat data set included only healthy forest, while the 1977 Landsat data set included all three forest classes. It seemed evident that moderately and heavily defoliated forest should be considered as change. However, previous studies concerned with gypsy moth defoliation have shown that moderately defoliated and healthy forest are not separable given the spectral and spatial resolutions of Landsat (Williams and Stauffer, 1978; Nelson, 1980; Williams and Ingram, 1981). Dottavio (1980) used a hand held radiometer to test the amount of light incident on the forest floor beneath these two stand conditions (healthy and moderately defoliated) and concluded that ". . . moderate defoliation (30-60% canopy removed) may have such high spectral variability that the class cannot be identified, even with higher spectral resolution, unless other environmental variables are considered." The evidence collected thus far indicate that only heavy defoliation can be consistently separated. Hence, it may be presumptuous to consider moderate defoliation as a changed area when dealing with Landsat data. This question was resolved by testing the difference and ratio images over a wide range of standard deviation threshold levels. The objective of the test was to determine whether or not moderately defoliated areas were alarmed as change more frequently than healthy forest at the various threshold levels. If the moderately defoliated pixels showed up as change significantly more often than the healthy pixels in either the difference or ratio images, then the two classes should be considered separately; i.e., moderate defoliation would be a valid change class that should not be grouped with the healthy forest. A paired-difference t test was run on each band of the difference and ratio images with standard deviation threshold levels (henceforth called

thresholds) ranging from 0.00 to 2.50, at intervals of 0.25 standard deviations (sd). Each transformed image was analyzed independently; if significant differences were found in either, then moderate defoliation would be considered a valid change class. The one sided statistical hypothesis follows:

H_0 : Moderate defoliation is alarmed as change less often or equally as often as healthy forest.

i.e., $U_d \leq 0$ when $U_d = U_{mod} - U_{healthy}$

H_1 : Moderate defoliation is alarmed more often than healthy forest.

i.e., $U_d \geq 0$

The tests showed that moderately defoliated areas were alarmed significantly more often than healthy forest areas in both the difference and ratio images. Hence it was concluded that the moderate defoliation class was a valid entity and was included as a change class when assessing accuracies.

The second question, that concerned with an accuracy criterion, deals with the characteristics of various accuracy measures and their interaction with the study site characteristics. This study was designed to define parameter levels for the technique which maximized the accuracy of delineating defoliated areas. In order to meet that objective, an appropriate accuracy measure had to be selected. Most researchers base their accuracy measurement on classification results tabulated in a confusion matrix. Two measures may be quickly calculated from such a matrix; (1) average accuracy, which weights each class equally and (2) overall accuracy, which weights each class according to the number of pixels in that class.

Only two classes are considered in the accuracy assessment, no change (healthy forest) and change (the two defoliated forest classes). Each of these accuracy criteria has an inherent bias due to the number of pixels involved

in each of the classes of interest. Table 2 details the class sizes.

Table 2. Number of pixels in each ground reference cover type class

<u>Change Class</u>	<u>Forest Class</u>	<u>Number of Pixels</u>
No Change	Healthy Forest	31067
Change	Moderate Defoliation	3307
	Heavy Defoliation	801
		} 4108
	Nonforest	<u>26887</u>
	Total	62062

Maximizing average accuracy would tend to favor low threshold levels. Average accuracy equally weights a class with 4108 pixels and a class with 31067 pixels. 3107 of those 4108 pixels are in a category which is usually confused with healthy forest. Hence as threshold levels increase, the classification accuracy of 75% ($\frac{3107}{4108}$) of the pixels in the change category quickly plummets. The higher thresholds, then, adversely affect average classification accuracy. Maximizing overall accuracy, on the other hand, would favor the high threshold levels which would correctly classify healthy forest at the expense of the defoliated classes.

To overcome the biases of the two accuracy criteria, a decision was made to maximize the average of the average accuracy and the overall accuracy. This combined accuracy figure, though in itself meaningless, does have the admirable characteristic of dampening the biases found in the first two criteria. The class weights associated with the combined accuracy criterion are merely the average of the weights of the average and overall accuracies, and are detailed in Table 3.

Table 3. Class weights for average, overall, and combined accuracy.

<u>Change Class</u>	<u>Forest Class</u>	<u>Average Accuracy</u>	<u>Class Weights</u> <u>Overall Accuracy</u>	<u>Combined Accuracy</u>
No Change	Healthy	0.50	0.8832 ¹	0.6916 ²
Change	Moderate Def.	0.4025 ³	0.0940 ⁵	0.3084 ⁷
	Heavy Def.	0.0975 ⁴	0.1168	
	Total	1.00	1.00	1.00

See Table 2 for sample sizes of each forest class.

- | | |
|----------------------------------|---------------------------------|
| 1. $31067/35175 = 0.8832$ | 5. $3307/35175 = 0.0940$ |
| 2. $(0.50 + 0.8832)/2 = 0.6916$ | 6. $801/35175 = 0.0228$ |
| 3. $(3307/4108) (0.50) = 0.4025$ | 7. $(0.50 + 0.1168)/2 = 0.3084$ |
| 4. $(801/4108) (0.50) = 0.0975$ | |

An alternate method which may have proved fairer would have been to arbitrarily assign a class weight of 0.50 for healthy forest, and 0.25 for the two defoliated cover types. Such an approach would have been more equitable for the defoliated classes, though an argument can be made for combined accuracy which slightly favors the class which comprises 88% of the test pixels. The combined accuracy figure (as set forth in Table 3) was used in the procedures outlined below in an attempt to find the method which best delineated gypsy moth defoliation.

D. Detecting Change Using Individual Bands

Each band of the difference and ratio images, as well as the single band of the difference of ratios image, were individually tested to determine the optimal threshold level for delineating defoliation. The "goodness" of a particular approach was judged solely by its quantitative performance; the higher the combined accuracy, the better the approach.

Greyscale cutoff values were calculated using threshold levels ranging from 0.00¹ to 2.50 standard deviations, using 0.25 sd increments. An output change image was produced for each threshold level. This change image

¹ Remember that only values on the low or high side of the mean were thresholded, depending on the band being processed. Hence the 0.00 standard deviation threshold level would result in roughly half the forested area being flagged as change.

contained zeros and ones, where "one" indicated change. This 0/1 change image was compared to the GRI and the combined accuracy was calculated. Hence for each individual band, there were 11 threshold/combined accuracy pairs.

The response surface around the threshold that produced the highest combined accuracy for a particular band was investigated further. For instance, if 1.00 standard deviation produced the highest combined accuracy for the band 6 difference image, then thresholds from 0.80 to 1.20 standard deviations (in increments of 0.05 sd) were tested. The threshold that produced the highest combined accuracy over this more limited range was considered the optimal threshold. In the event that a range of thresholds produced the same maximum combined accuracy (as was typically the case with the difference image), then the mid point of the range was considered optimal. The optimal threshold for each band of the difference and ratio images served as the starting point or "best guess" threshold for the multiband image analysis.

E. Modeling Combined Accuracy - Single Band Images

Identification of the optimal threshold level involved time which was, perhaps, unnecessarily spent. If the sd threshold/combined accuracy relationship could be accurately modeled using the initial empirical results (i.e., thresholds ranging from 0.0-2.5 sd, increments of 0.25), then the more detailed investigation (every 0.05 sd) could be bypassed. Though finding such a relationship would not help this research (since the 0.05 sd investigation would have been completed), future work might benefit. Hence, the threshold/combined accuracy relationship was modeled using data from the initial response surface analysis (0.00-2.50 sd, 0.25 increments). The model derived from this data was used to calculate the expected combined accuracy values around the initial empirical maximum. These expected values were compared to the actual accuracy values.

One would expect that a certain threshold would produce the maximum accuracy, and the accuracies would become smaller as the threshold moved toward 0.00 or 2.50 standard deviations. Hence, a second order (quadratic) polynomial was fit to the threshold/combined accuracy data. It was expected that the resultant curves would be concave downward (2nd derivative negative) and that the solution to the first derivative set equal to zero would be the threshold that would maximize the combined accuracy.

A second response surface modeling approach was tested as a means of interpolating between the empirical results to try and find the threshold that would yield the maximum combined accuracy for a particular band. The threshold and accuracy data for each band was modeled using a natural cubic spline function. The spline function fits a third order polynomial between each pair of knots, or data coordinates (each threshold/accuracy observation). Hence, the empirical information is modeled perfectly; the spline may be evaluated between the knots for purposes of interpolation. An explanation of the approach and an example of the Fortran program similar to the one used in this study may be found in Forsythe, Malcolm and Moler (1977). The interpolated accuracies were compared to the actual accuracies to see which, if either, model produced acceptable results.

F. Multiband Analysis Difference and Ratio Images

A sequential simplex design was used to explore the combined accuracy response surface for various band combinations. The purpose of this portion of the experiment was to explore the response surface of these multiband images to maximize the accuracy for any particular band combinations. These maximized accuracies were compared to determine (1) if additional bands provided additional information and (2) which band combination best characterized forest canopy alteration due to gypsy moth.

A simplex design is one in which the experimenter mathematically brackets a "guessed at" maximum with different treatment combinations so that the response surface around that maximum can be explored. "Sequential" refers to the fact that the outcomes of previous treatment combinations are used to establish new treatment combinations. The experimental approach is such that the experimenter steps across the response surface to the maximum yield or response. The logic and mathematics of the approach may be found in Anderson and McClean, 1974, pg, 362-367; these pages are reproduced in Appendix A.

The best two, three and all four bands for the difference and ratio images were considered in the multiband analysis. The individual band ratings for the difference or ratio image were based on the best empirical performance of each band. A composite change image was produced for a given band combination by adding the change/no change results for the individual bands (see Figure 8). A pixel was considered changed if any of the bands found it changed.

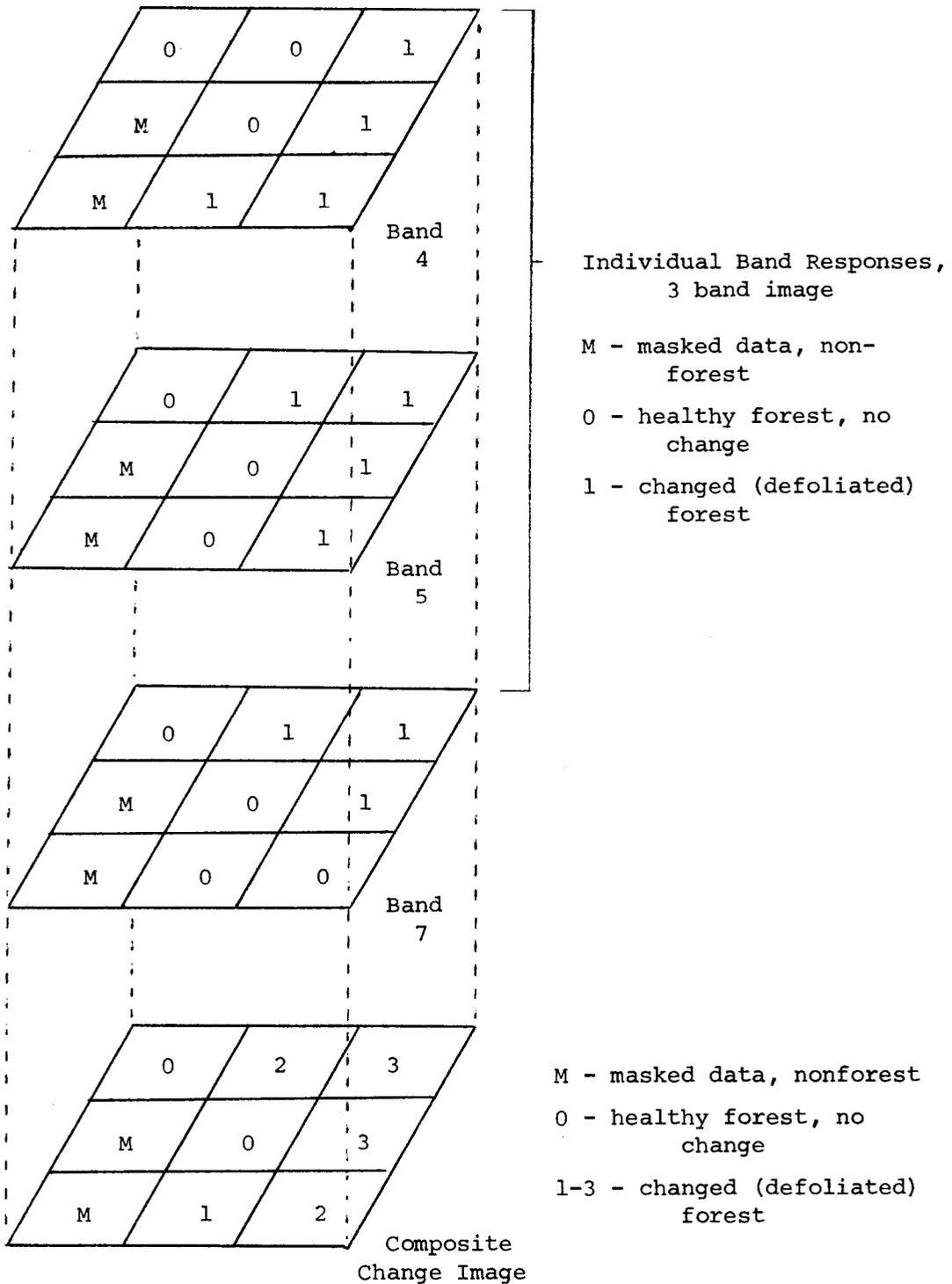


Figure 8 Construction of a change image - multiband analysis.

A pixel is considered changed if any of the bands involved note change.
 The composite change image is compared to the ground reference image.

V. RESULTS

A. Analysis of Individual Bands

Individual bands of the various transformed images were investigated to determine which bands best classified defoliated areas as change. The following facts should be considered while reading this section:

1. Combined accuracy was calculated by averaging overall and average classification accuracies. The latter two were calculated from performance statistics of the individual forest classes, the combined accuracy figure was the criterion maximized throughout the study.

2. The study is entirely quantitative. The sole measure of the ability of a given technique to accurately classify defoliation as change is the combined accuracy figure. No visual or qualitative judgements were involved in the selection procedure.

3. No sampling error was involved in these accuracy assessments; the assessment included all of the forested pixels as defined by the forest/non-forest mask.

4. Non-forest areas were masked and any changes in the non-forest areas were not considered in the accuracy figures.

5. The change images output for single bands were 0/1 masks. A pixel was correctly classified if a 0 on the change image corresponded to healthy forest on the GRI or if a 1 on the change image corresponded to moderate or heavy defoliation on the GRI.

6. Only one side of the population distribution of a given band was considered for detecting alterations due to gypsy moth defoliation. Tests were done where outliers on both sides of the mean were considered over a range of thresholds. The results from these tests were compared to the results when only one side was considered. In all bands tested (the 4 difference and

4 ratio bands) the highest accuracy of the one-sided approach was higher than the highest accuracy when both sides of the mean were considered.

1. Difference Image:

a. Empirical Results

Table 4a-d presents the results of the threshold tests. The accuracy of classification of the individual cover types and the average, overall, and combined accuracies are presented for the various thresholds. The two visible channels (4 and 5) were analyzed to the low side of the mean, vice versa for the infrared channels. Since the expected spectral shift due to gypsy moth defoliation would be in these directions.

Note that as the sd level increases, the healthy forest accuracy, hence the overall accuracy, increases. Conversely, the defoliated cover type class accuracies and (with one exception) average accuracy fall. The combined accuracy describes a concave downward curve with an empirical maximum of 1.00 and 1.25 for band 4, 0.50 for bands 5 and 7 and 1.00 sd for band 6. The channel best suited for discriminating levels of defoliation from healthy forest was band 5, band 7 was second best, and band 6 was the worst.

Confusion matrices may be found in Appendix B with the methods used to calculate the various accuracy measures.

Table 4. Threshold level vs. classification accuracy - difference image.

SD Level: standard deviation threshold which produced these accuracies
 Hlthy: % correct classification, healthy forest
 Mod: % correct classification, moderate defoliation
 Hvy: % correct classification, heavy defoliation
 Chng: % correct classification, change areas. Those areas noted as defoliated on the GRI were considered Change. Hence, Chng = (# of change pixels in moderate defoliation + # of change pixels in heavy defoliation) ÷ (total number of mod and hvy pixels in GRI).
 Avg: Average classification accuracy = $\frac{\% \text{ correct chng} + \% \text{ correct Nochange}}{2}$, note that % correct nochange and % correct healthy are equivalent.
 Over: Overall classification accuracy = total # correctly classified pixels/total # of pixels.
 Comb: Combined classification accuracy: (Avg + Over)/2.

a) Difference Image, Band 4

SD Level	% Correct Classification						
	Hlthy	Mod	Hvy	Chng	Avg	Over	Comb
0.0	38.97	87.30	99.38	89.65	64.31	44.89	54.60
0.25	38.97	87.30	99.38	89.65	64.31	44.89	54.60
0.50	69.59	62.72	91.39	68.31	68.95	69.44	69.20
0.75	69.59	62.72	91.39	68.31	68.95	69.44	69.20
1.00	89.32	29.21	71.66	37.49	63.41	83.27	73.34
1.25	89.32	29.21	71.66	37.49	63.41	83.27	73.34
1.50	96.34	8.65	46.57	16.04	56.19	86.96	71.57
1.75	98.49	1.51	19.98	5.11	51.80	87.58	69.69
2.00	98.49	1.51	19.98	5.11	51.80	87.58	69.69
2.25	99.27	0.03	6.87	1.36	50.31	87.83	69.07
2.50	99.27	0.03	6.87	1.36	50.31	87.83	69.07

b) Difference Image, Band 5

SD Level	% Correct Classification						
	Hlthy	Mod	Hvy	Chng	Avg	Over	Comb
0.0	65.81	74.24	98.63	78.99	72.40	67.35	69.87
0.25	65.81	74.24	98.63	78.99	72.40	67.35	69.87
0.50	83.40	49.89	97.13	59.10	71.25	80.56	75.91
0.75	90.92	27.85	91.39	40.24	65.58	85.00	75.29
1.00	94.48	11.37	81.02	24.95	59.72	86.36	73.04
1.25	96.38	5.65	66.67	17.55	56.97	87.18	72.07
1.50	97.62	2.69	57.30	13.34	55.48	87.78	71.63
1.75	98.33	1.15	45.19	9.74	54.03	87.98	71.01
2.00	98.80	0.39	32.83	6.72	52.76	88.05	70.40
2.25	99.08	0.12	23.47	4.67	51.92	88.14	70.03
2.50	99.38	0.12	15.48	3.12	51.25	88.14	69.69

Table 4 (continued):

c) Difference Image, Band 6

% Correct Classification							
<u>SD Level</u>	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Over</u>	<u>Comb</u>
0.0	72.28	39.01	93.01	49.54	60.91	69.62	65.27
0.25	79.79	29.51	91.39	41.58	60.69	75.33	68.01
0.50	89.37	15.78	85.14	29.31	59.34	82.35	70.85
0.75	92.17	12.43	81.02	25.80	58.99	84.42	71.70
1.00	95.59	7.35	72.66	20.08	57.84	86.77	72.31
1.25	96.53	5.68	66.79	17.60	57.06	87.31	72.19
1.50	97.86	2.87	53.06	12.66	55.26	87.91	71.58
1.75	98.60	1.18	42.07	9.15	53.88	88.15	71.01
2.00	98.90	0.79	35.46	7.55	53.22	88.23	70.73
2.25	99.25	0.33	21.60	4.48	51.86	88.18	70.02
2.50	99.36	0.15	15.86	3.21	51.28	88.13	69.71

d) Difference Image Band 7

% Correct Classification							
<u>SD Level</u>	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg.</u>	<u>Over</u>	<u>Comb</u>
0.0	75.27	54.85	98.13	63.29	69.28	73.87	71.58
0.25	84.15	40.37	96.00	51.22	67.69	80.31	74.00
0.50	89.75	30.06	93.01	42.33	66.04	84.21	75.13
0.75	93.21	19.93	89.51	33.50	63.35	86.24	74.80
1.00	95.19	13.61	84.39	27.41	61.30	87.28	74.29
1.25	96.56	9.71	79.03	23.22	59.89	88.00	73.94
1.50	97.48	6.32	71.79	19.08	58.28	88.32	73.30
1.75	98.07	4.11	63.30	15.65	56.91	88.54	72.73
2.00	98.64	2.36	51.44	11.93	55.28	88.51	71.90
2.25	99.00	1.09	40.32	8.74	53.87	88.45	71.16
2.50	99.22	0.45	30.09	6.23	52.73	88.36	70.54

A second, limited, response surface analysis was performed around the threshold level which produced the highest accuracy for each band in Table 4. Hence, threshold levels from 0.80 sd to 1.45 sd (in 0.05 sd increments) were investigated for band 4 because this range brackets the thresholds (i.e., 1.00, 1.25) which yielded the highest combined accuracy. The highest combined

accuracy in bands 5 and 7 were associated with the 0.50 standard deviation threshold, hence thresholds from 0.30-0.70 standard deviations were investigated, in 0.05 intervals. The region from 0.80-1.20 standard deviations was investigated in band 6. The results of this refined empirical analysis are given in Table 5 and in Figure 9a-d. The best empirical results as noted in Table 5 provided input to the multiband analysis.

b. Modeling - Difference Image

The results in Table 4 (threshold vs. combined accuracy) were modeled using (1) a quadratic regression equation and (2) a natural cubic spline function. The models were evaluated to determine which most closely fit the empirical data. Figure 9 a-d provides a graphical comparison, while Table 5 provides a tabular comparison of the modeled and empirical results.

Table 5. Difference Image, Empirical vs. Modeled results.

Band	Best Empirical sd levels	Best Empirical Comb. Acc (%)	Modeled Optimal Results				
			Quadratic			Spline*	
			sd thresh	Pred. Acc	R ²	sd thresh	Pred. Acc
4	0.85-1.25	73.34	1.56	73.13	0.81	1.10	73.99
5	0.35-0.60	75.91	1.02	73.15	0.45	0.60	76.41
6	0.95-1.05	72.31	1.42	72.28	0.87	1.10	72.36
7	0.45-0.65	75.13	0.94	74.35	0.83	0.55	75.15

* The spline functions models the data perfectly, the derived functions are used for interpolation only. Hence an R² value is not applicable; if calculated it would always be 1.00

In Table 5 and in Figure 9 a-d, the concern lies with how well a particular modeling approach estimates the optimal sd threshold. One notes that the spline function seems to do well in this regard; three of the four optimal thresholds predicted by the spline are within the limits of the empirical results. The quadratic model does a very poor job of predicting the optimal threshold level, none are reasonably close.

Figure 9 a
Difference image,
band 4.

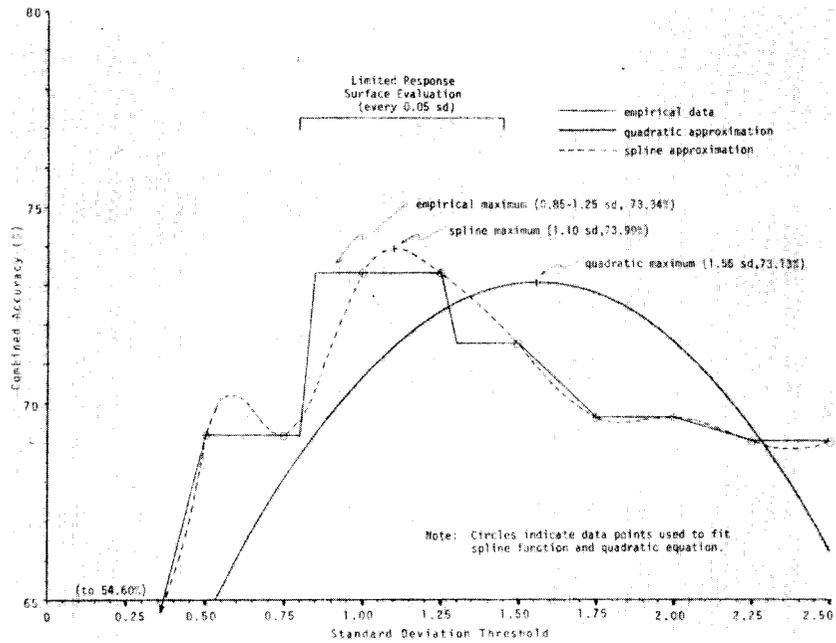


Figure 9 b
Difference image,
band 5.

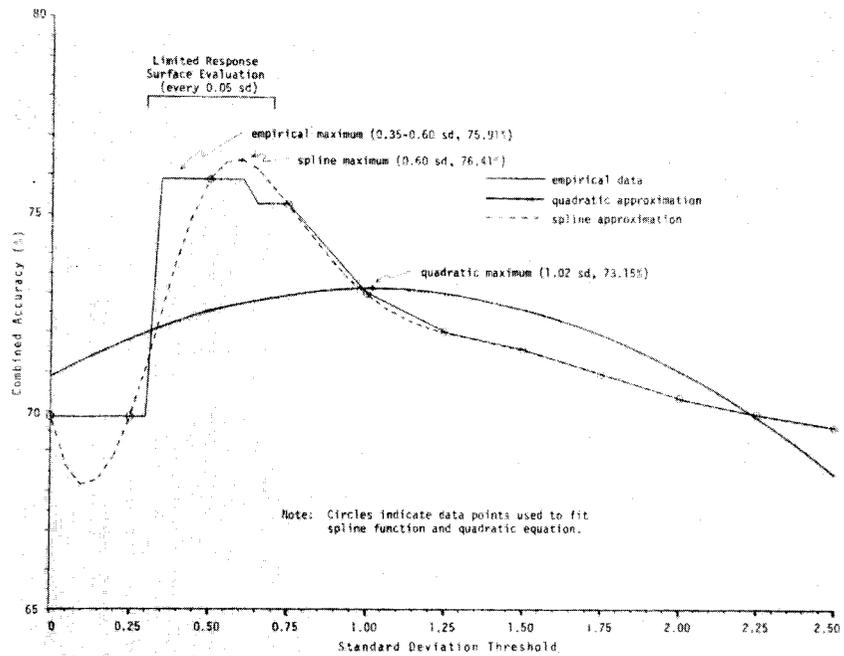


Figure 9 a-d Modeling the threshold/combined accuracy relationship using a quadratic equation and a cubic spline function, difference image, all four bands.

Figure 9 c
Difference image,
band 6.

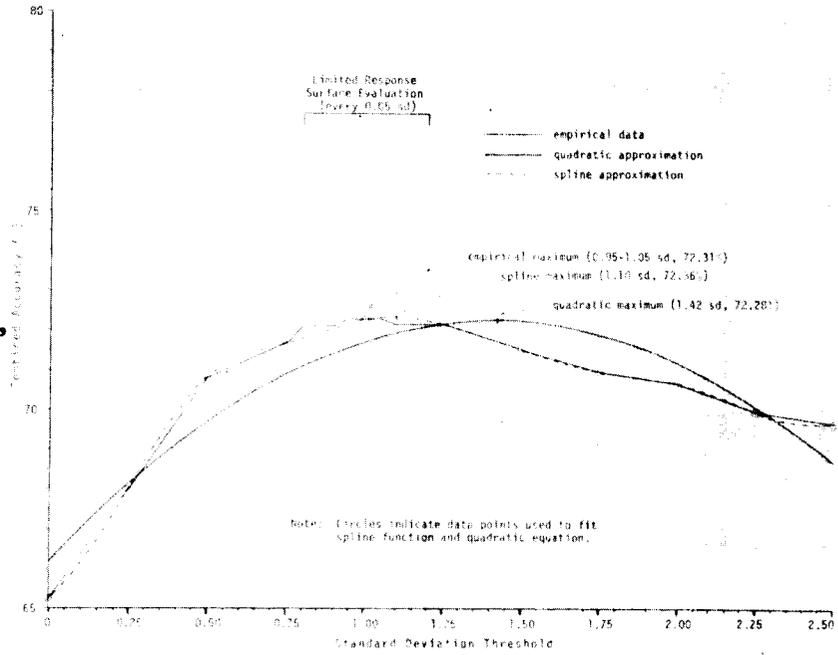
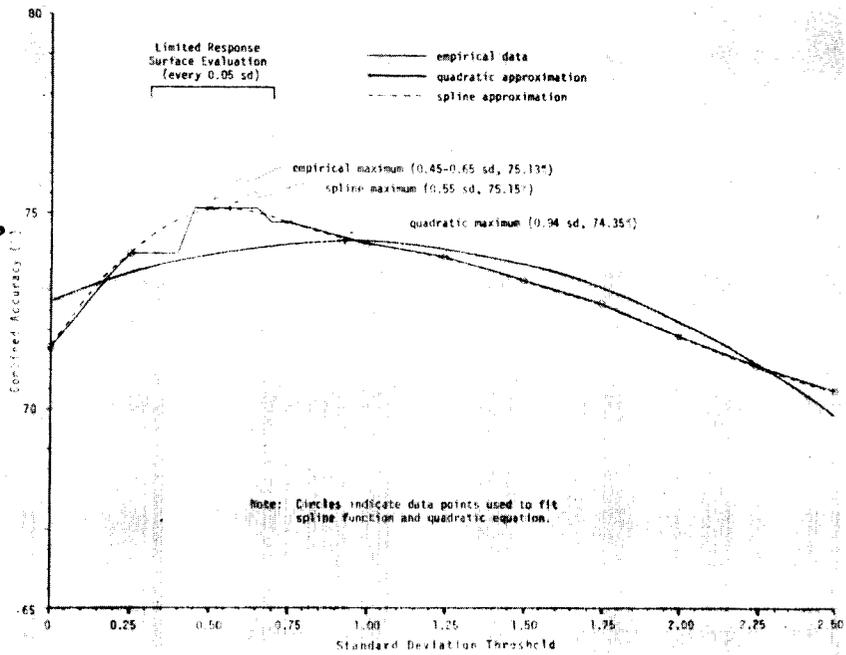


Figure 9 d
Difference image,
band 7.



2. Ratio Image:

a. Empirical Results

Tables 6 a-d present the results of the threshold tests for the ratio image. The ratio image channels were handled in an inverse manner relative to the difference image channels. The two visible ratio bands were analyzed to the high side of the mean only, the infrared channels to the low side. These areas of the channel populations are those where one would expect to pick up the gypsy moth defoliation due to the nature of the calculations producing the ratio image ($\frac{\text{Band X, 1977}}{\text{Band X, 1976}}$) and due to the spectral shifts caused by foliage removal.

Table 6. Threshold level vs. classification accuracy, Ratio image.

NOTE: See Table 4 for explanation of column headings

a) Ratio Image, Band 4

SD Level	% Correct Classification						
	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Over</u>	<u>Comb</u>
0.0	39.97	87.30	99.38	89.65	64.31	44.89	54.60
0.25	69.59	62.72	91.39	68.31	68.95	69.44	69.20
0.50	69.59	62.72	91.39	68.31	68.95	69.44	69.20
0.75	88.80	29.24	71.66	37.51	63.16	82.81	72.98
1.00	89.36	29.21	71.66	37.49	63.43	83.31	73.37
1.25	91.32	29.00	67.92	36.59	63.95	84.93	74.44
1.50	96.44	8.65	46.57	16.04	56.24	87.05	71.65
1.75	97.33	8.56	44.94	15.65	56.49	87.79	72.14
2.00	98.23	2.96	20.35	6.35	52.29	87.50	69.89
2.25	98.89	1.51	19.98	5.11	52.00	87.93	69.97
2.50	99.20	1.42	15.73	4.21	51.70	88.10	69.90

b) Ratio Image, Band 5

SD Level	% Correct Classification						
	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Over</u>	<u>Comb</u>
0.0	65.81	74.24	98.63	78.99	72.40	67.35	69.87
0.25	69.34	72.12	98.25	77.21	73.28	70.26	71.77
0.50	84.00	49.89	97.13	59.10	71.55	81.09	76.32
0.75	90.29	28.39	91.39	40.68	65.48	84.49	74.99
1.00	92.75	27.64	91.01	40.00	66.37	86.59	76.48
1.25	94.91	12.37	81.40	25.83	60.37	86.85	73.61
1.50	96.35	9.53	74.53	22.20	59.28	87.69	73.48
1.75	97.37	6.38	66.29	18.06	57.71	88.11	72.91
2.00	97.89	4.17	61.42	15.34	56.61	88.25	72.43
2.25	98.41	2.69	57.43	13.36	55.89	88.47	72.18
2.50	98.73	2.15	50.31	11.54	55.14	88.55	71.84

c) Ratio Image, Band 6

% Correct Classification

<u>SD Level</u>	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Over</u>	<u>Comb</u>
0.0	72.28	39.01	93.01	49.54	60.91	69.62	65.26
0.25	85.06	22.53	88.39	35.37	60.22	79.26	69.74
0.50	91.99	12.82	82.02	26.31	59.15	84.32	71.73
0.75	95.45	7.95	75.53	21.13	58.29	86.77	72.53
1.00	97.14	4.78	64.54	16.43	56.79	87.72	72.25
1.25	98.30	2.57	52.81	12.37	55.33	88.26	71.80
1.50	98.89	1.06	42.57	9.15	54.02	88.41	71.21
1.75	99.33	0.51	32.33	6.72	53.02	88.51	70.77
2.00	99.56	0.12	19.10	3.82	51.69	88.38	70.04
2.25	99.69	0.03	8.86	1.75	50.72	88.26	69.49
2.50	99.78	0.0	1.75	0.34	50.06	88.17	69.12

d) Ratio Image, Band 7

% Correct Classification

<u>SD Level</u>	<u>Hlthy</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Over</u>	<u>Comb</u>
0.0	60.95	70.97	98.88	76.41	68.68	62.75	65.72
0.25	91.28	25.58	92.63	38.66	64.97	85.14	75.05
0.50	97.05	8.32	76.53	21.62	59.33	88.24	73.79
0.75	98.96	1.42	45.07	9.93	54.45	88.56	71.50
1.00	99.63	0.06	13.61	2.70	51.17	88.31	69.74
1.25	99.88	0.0	0.0	0.0	49.94	88.22	69.08
1.50	99.94	0.0	0.0	0.0	49.97	88.27	69.12
1.75	99.98	0.0	0.0	0.0	49.99	88.30	69.15
2.00	99.99	0.0	0.0	0.0	50.00	88.32	69.16
2.25	100.00	0.0	0.0	0.0	50.00	88.32	69.16
2.50	100.00	0.0	0.0	0.0	50.00	88.32	69.16

The order of band usefulness for delineating gypsy moth defoliation is identical to that of the difference image, best to worst: bands 5, 7, 4, 6. Each ratio band was subject to further investigation around the threshold which produced the highest combined accuracy in Table 6. Band 4 was checked from 1.05 to 1.45 sd (0.05 increments); Band 5, 0.80-1.20; Band 6, 0.55-0.95; Band 7, 0.05-0.45. The results of this refined empirical analysis are given in Table 7 and in Figure 10a-d. The best empirical results, as noted in Table 7, provided input to the multiband analysis.

Figure 10 a
Ratio image,
band 4.

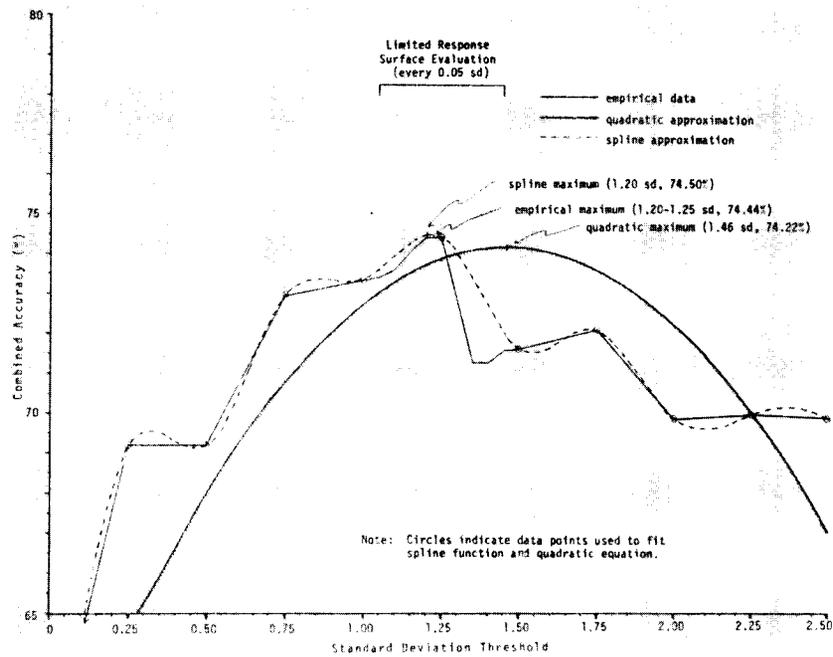


Figure 10 b
Ratio image,
band 5.

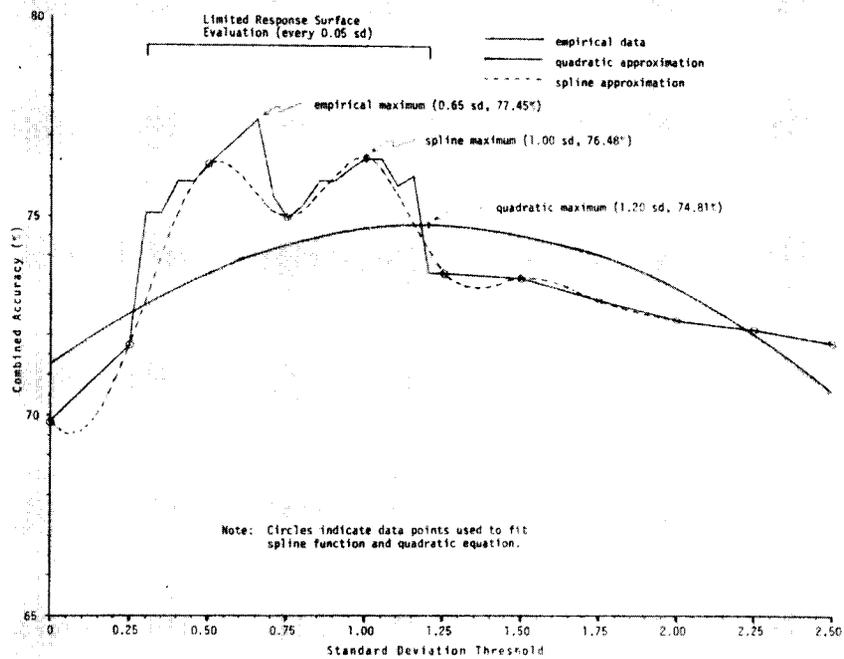


Figure 10 a-d Modeling the threshold/combined accuracy relationship using a quadratic equation and a cubic spline function, ratio image, all four bands.

Figure 10 c
Ratio image,
band 6.

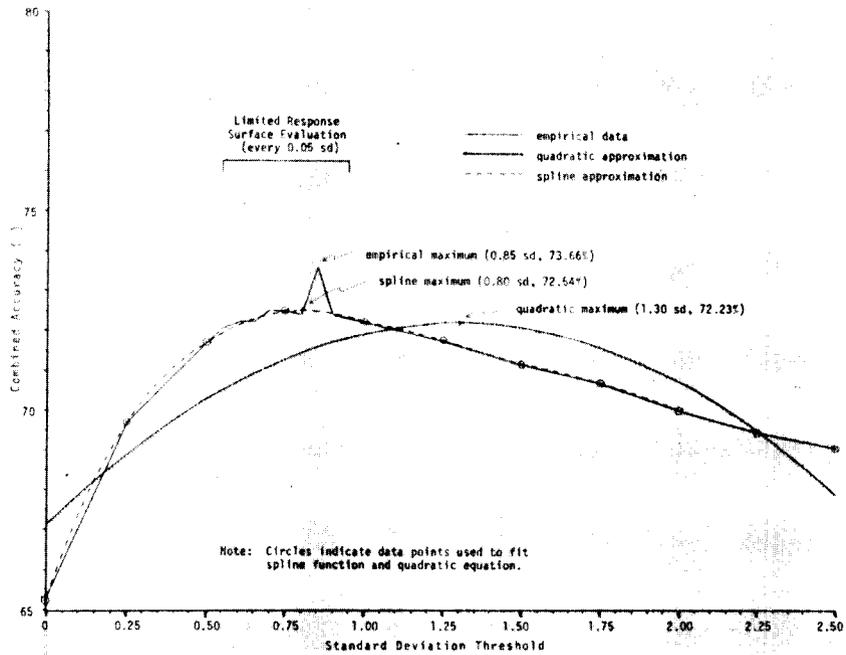
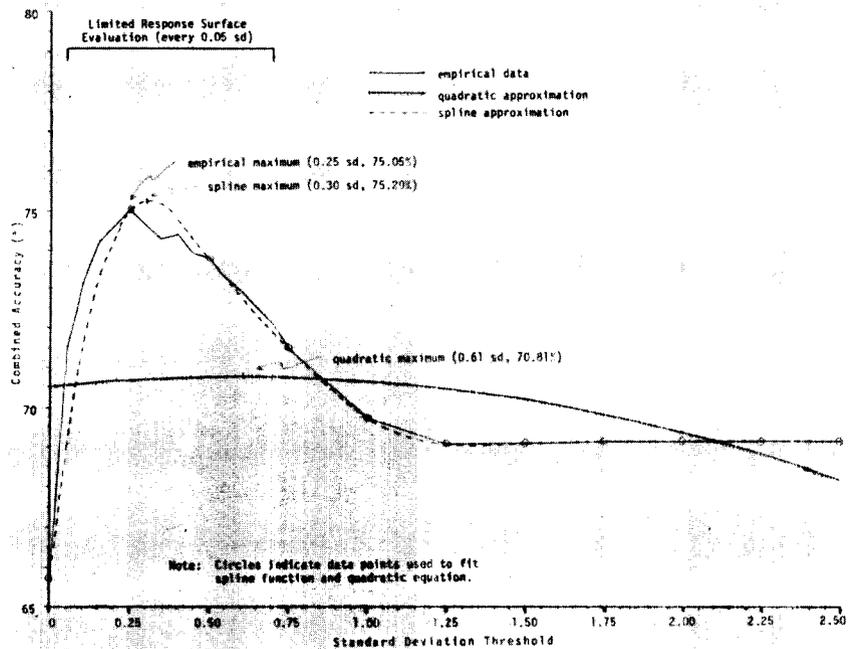


Figure 10 d
Ratio image,
band 7.



b. Modeling - Ratio Image:

The quadratic regression model and the natural cubic spline model were used to analyze the threshold/combined accuracy data in Table 6a-d. The model results were compared to the empirical results at the points mentioned in the paragraph above. Table 7 and Figures 10a-d summarize the results.

Table 7. Ratio Image, Empirical vs. Modeled Results

Band	Best Empirical sd level	Best Empirical Comb. Acc (%)	Modeled Optimal Results				
			Quadratic		R ²	Spline	
			sd thresh	Pred. Acc			sd thresh
4	1.20-1.25	74.44	1.46	74.22	0.72	1.20	74.50
5	0.65	77.45	1.20	74.81	0.52	1.00	76.48
6	0.85	73.66	1.31	72.23	0.73	0.80	72.54
7	0.25	75.05	0.61	70.81	0.12	0.30	75.29

Again, the spline model performed well, the quadratic model did very poorly. The spline consistently predicted the optimal standard deviation within 0.05 with the exception of band 5. The extreme irregularity of the threshold/accuracy relationship in this band produced the inaccurate spline response (see Figure 10b).

3. Difference of Ratios Image

a. Empirical Results

This image was evaluated to determine if a one band comparison of vegetation indices at time 1 and time 2 would produce more accurate results than the albedo comparisons of the Ratio and Difference images. The 7/5 ratio was calculated for each pixel in the 1976 (healthy forest) data set and for each pixel in the 1977 (defoliated) data set. The 1977 values were subtracted

from the 1976 ratios and the mean and standard deviation of the new image was calculated. Since the 7/5 ratio in defoliated areas is expected to decrease, only values at the high end of the scale (greater than the mean) were considered to potentially be gypsy moth defoliation. The results of the empirical analysis are given in Table 8.

Table 8. Threshold level versus classification accuracy, Difference of Ratios image.

SD Level	% Correct Classification						
	Hlthy	Mod	Hvy	Chng	Avg	Over	Comb
0.00	64.00	74.47	98.88	81.65	72.85	66.11	69.48
0.50	80.26	58.87	98.50	66.60	73.43	78.67	76.05
1.75	85.87	49.44	97.75	58.86	72.36	82.71	77.54
1.00	89.79	40.64	96.75	51.58	70.68	85.32	78.01
1.25	92.58	32.45	94.01	44.45	68.51	86.96	77.74
1.50	94.78	25.18	91.63	38.14	66.46	88.16	77.31
1.75	96.23	17.30	87.89	31.06	63.64	88.62	76.13
2.00	97.34	12.85	83.40	26.61	61.98	89.08	75.53
2.25	98.30	9.13	77.03	22.37	60.34	89.43	74.88
2.50	98.85	6.23	69.16	18.50	58.67	89.46	74.07

The response surface around 1.00 standard deviations (from 0.80-1.20) was checked to see if higher accuracies could be obtained. The 1.00 standard deviation level yielded the only combined accuracy over 78%. On a single band basis, this technique performed slightly better than the best band of the differenced or ratioed data.

b. Modeling - Difference of Ratios

A quadratic regression expression and a natural cubic spline function were fit to the data in Table 8. Table 9 and Figure 11 summarize the findings of these analyses.

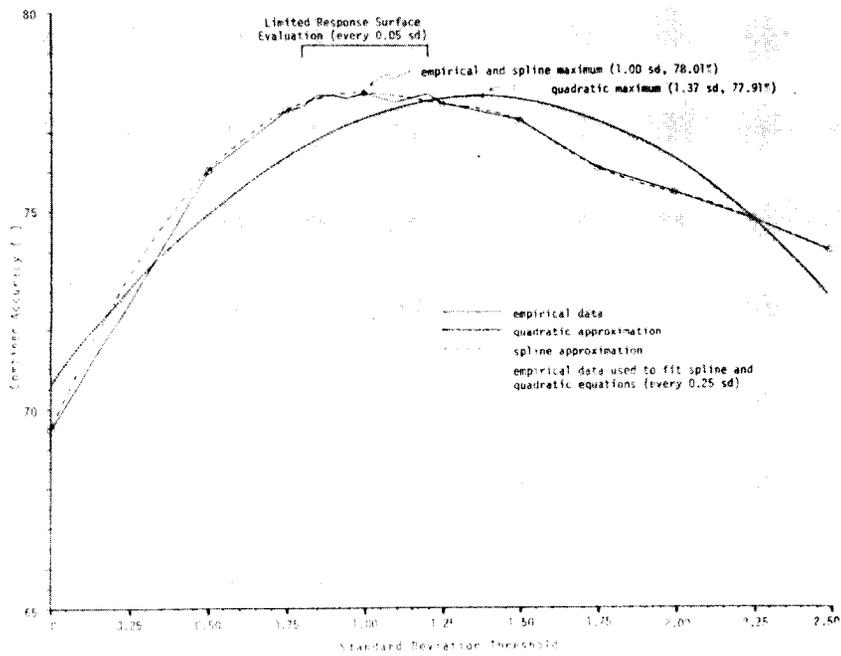


Figure 11 Modeling the threshold/combined accuracy relationship using a quadratic equation and a spline function, difference of ratios image.

Table 9. Difference of Ratios - Empirical vs. Modeled Results

	Best Empirical sd level	Best Empirical Comb. Acc (%)	Modeled Optimal Results				
			Quadratic			Spline	
			sd thresh	Pred. Acc	R ²	sd thresh	Pred. Acc
Difference of Ratios	1.00	78.01	1.37	77.91	0.86	1.00	78.01

The quadratic equation overestimated the optimal threshold; the spline function accurately modeled the response surface. Thus, the spline function predicted the optimal threshold within 0.05 standard deviations in eight of nine trials. The quadratic did poorly in all nine.

B. Multiband Image Analysis

Having found the optimal sd levels for individual bands in differenced or ratioed images, the task became one of determining whether or not information from two or more bands improved change detection capability. The single band analysis highlights the fact that optimal threshold levels are not the same for each band. For instance, the optimal threshold for band 4 is twice the optimal threshold for band 5 in the difference image. Hence, the levels for the individual bands used in combination must be adjusted independently when the response surface of two or more bands is studied.

A full factorial design would be the most informative approach to determine the optimal threshold levels of two or more bands used in combination. The drawback to this approach is the number of treatment combinations necessary to adequately study the entire response surface. For instance, if 11 standard deviation levels are to be tested using all 4 channels to detect change, the number of treatment combinations that must be tested is 11^4 or 14,641. Full factorials, though ideal in the context of information content, were not run due to time and expense considerations.

A sequential simplex experimental design was used to investigate the band combination problem. Sequential refers to the fact that the results from

previous trials are used to formulate the new treatment levels which are subsequently tested. A simplex design is one in which a treatment combination of interest is bracketed with experimental treatment combinations. A simplex design is a mathematical method of defining treatment level combinations in an area of the response surface that is of interest to the researcher. That area, or treatment combination of interest, might be a "guessed at" maximum, for instance, a treatment combination in a chemical experiment that produced the maximum yield of a desired product. The sequential simplex design permits the experimenter to study the response surface and adjust factor levels such that, eventually, he finds the optimal treatment combination that maximizes the dependent variable. An explanation of the approach and statistical "how-to" may be found in Anderson and McLean (1974); pg. 362-367. These pages are reproduced in Appendix A, with permission of the publisher.

The "best guess" initial estimates were those sd thresholds which yielded the highest combined accuracy for a particular band. In the event a range of thresholds produced the same high accuracy, the mid-point of the range was used. The initial thresholds for the multiband analysis are listed in Tables 5 and 7, second column.

The best two bands (5 and 7), the best three bands (5, 7, 4) and all four bands were tested to see whether the information supplied by the additional bands improved classification accuracy. To review, a multiband change image was produced by adding together the 0/1 change masks of the individual bands. For instance, the final change image (single band) of a three band (5, 7 and 4) treatment combination would have a zero where no band found change, a 1 where only one of the three found change, a two where two of the three bands found change, and a three where all bands concurred (see Figure 8 and associated text).

The results of the sequential simplex experiment are given in Table 10. The thresholds listed are, theoretically, the treatment levels that will yield the maximum response (highest combined accuracy) for a particular band combination. In actuality, these treatment levels may be viewed as approximations (i.e., close to the maximum) to the optimal treatment levels. Previous work showed that small improvements in classification accuracy (up to one half a percentage point) could be expected if a micro-response surface investigation around the simplex maximum was conducted.

Table 10: Results of the sequential simplex experiment to determine optimal threshold levels for bands used in combination.

Image	Channel Comb.	sd Threshold Level				% Correct Classification					
		Ch2	Ch4	Ch1	Ch3	Hth	Mod	Hvy	Avg	Over	Comb.
Dif.	5,7	0.47	1.70			83.05	49.89	97.13	71.08	80.26	75.67
Dif.	5,7,4	0.49	1.05	1.40		81.58	52.46	97.63	71.42	79.21	75.32
Dif.	5,7,4,6	0.41	0.95	1.33	1.22	81.21	52.46	97.63	71.24	78.88	75.06
Ratio	5,7	0.65	0.54			85.21	49.53	97.13	72.01	82.13	77.07
Ratio	5,7,4	0.95	0.93	1.19		86.71	44.60	94.38	70.51	82.93	76.72
Ratio	5,7,4,6	0.59	0.65	1.50	1.07	83.43	52.04	97.25	72.15	80.80	76.47

C. Comparison of All Approaches

Comparison of the results of the single and multiband analyses for the three images show that: (1) the difference of ratios transformation produced the highest combined classification accuracy; (2) the red channel, band 5, classifies gypsy moth defoliation more accurately than any other single channel or channel combination; (3) additional bands do not improve the discriminatory capability of differenced or ratioed data; (4) ratioed data provided consistently more accurate results than did the differenced data. Table 11 describes the best performance for each of the three images.

Table 11. Best classification performance for Difference, Ratio and Difference of Ratio images.

<u>Image</u>	<u>Band(s)</u>	<u>sd</u> <u>Thresh</u>	<u>Hth</u>	<u>Mod</u>	<u>Hvy</u>	<u>Chng</u>	<u>Avg</u>	<u>Overall</u>	<u>Combined</u>
Difference	5	0.48	83.40	49.89	97.13	59.10	71.25	80.56	75.91
Ratio	5	0.65	85.86	49.38	96.63	58.59	72.23	82.68	77.45
D of R	-	1.00	89.79	40.64	96.75	51.58	70.68	85.32	78.01

NOTE: The contingency tables associated with the best approaches are given in Appendix B.

Figure 12 presents each of the images listed in Table 10 and the ground reference image for visual comparison.

The results point to the fact that a Difference of Ratios transformation is the most useful method for detecting gypsy moth defoliation. Not only did it turn in the highest classification accuracy over a relatively wide range of threshold levels but it did so using only one band of transformed data. The ratio band 5 data performed almost as well, but did so over a more limited range of thresholds (compare Figures 10b and 11). Additional bands did not enhance the capability of a particular data transformation to detect change. The difference image universally turned in the poorest detection performances; an argument may be made for processing continuous (real) images instead of discrete (byte) images when thresholding to detect change.

D. Characteristics of Landsat Data and Transformations

A notable feature of the results given in the preceding sections is the strikingly low thresholds necessary to accurately classify change due to gypsy moth defoliation. Previous work had shown that threshold levels on the order of two to three standard deviations were typical (Stauffer and McKinney, 1978; Ingram et al., 1981). The low thresholds encountered indicated that the transformed populations were most likely markedly non-normal. Hence, a number of statistical tests were run on many of the data sets which were used in this study. The

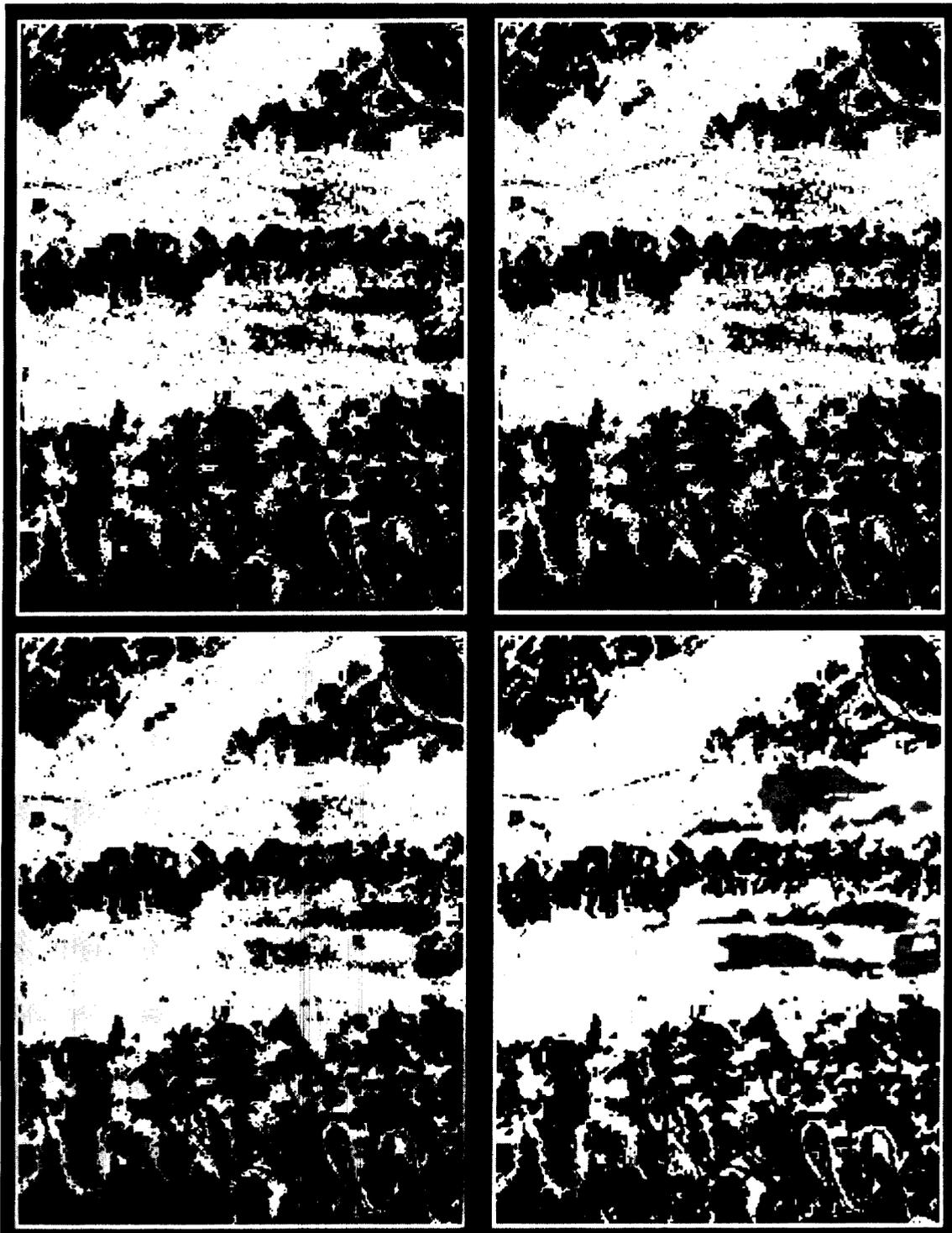


Figure 12 The best difference, ratio, and difference of ratios images.

Upper left: difference image, band 5	black - nonforest
Upper right: ratio image, band 5	white - healthy forest
Lower left: difference of ratios image	grey - changed (defoliated) forest
Lower right: ground reference image	black - nonforest
	white - healthy forest
	light grey - moderate defoliation
	dark grey - heavy defoliation

statistical tests were conducted to determine if a particular data set was normally distributed and to characterize the "shape" of the frequency distribution.

The raw Landsat data and the transformed data were tested using the SAS program Univariate (Statistical Analysis System, SAS Institute, Inc., 1979). Table 11A and B list the results of those tests which included the calculations of the mean, standard deviation, Fisher's g_1 and g_2 statistics (skewness and kurtosis), and the Kolmogorov-Smirnov D-statistic test for normality. These numbers allow the population to be characterized as follows:

1. Significantly nonnormal (Kolmogorov-Smirnov Test);
2. Significant skew, left or right (Fisher's g_1 statistic);
3. Significant kurtosis, leptokurtic or platykurtic (Fisher's g_2 statistic).

The Kolmogorov-Smirnov statistic is a nonparametric value which deals with the maximum deviation of the observed population from the theoretical (or normal) population (Davis, 1973, pg. 276). The Fisher g statistics allow one to draw conclusions concerning the shape of the population in question. The g_1 and g_2 statistics have an expected value of zero in a normally distributed population and are themselves normally distributed with a variance of $6/N$ and $24/N$ (these are approximations, see Fisher, 1970, pg. 75 for precise variance calculations), where N is the number of values sampled from the population.

In this case, 252 points were systematically sampled over the entire 286 line by 217 column image, corresponding to every 16th pixel in a line (14 sample points per line), every 15th line (18 lines). The 252 sample points represented the range of cover types found on the Wertzville Quad. The forested area, which was the sole consideration in the change detection study, was analyzed separately by multiplying the data by a 0/1 forest mask. Only 146 sample points remained after masking, indicating that roughly 57.9% of the

quadrangle is forested (the actual forested area defined by the forest/non-forest mask is 56.7%). The following data was tested for normality:

1976 "raw" Landsat data - 4 channels

1977 "raw" Landsat data - 4 channels

Difference data - 4 channels (DIF)

Ratioed data - 4 channels (RAT)

Band7/Band5 ratio, 1976 data - 1 channel (7/5, 1976)

Band7/Band5 ratio, 1977 data - 1 channel (7/5, 1977)

Difference of Ratios - 1 channel (DR)

All sampled points in each band in a particular data set were analyzed (252 points, Table 11A), then just the forested points were tested (146 points, Table 11B).

A data set that is normal will have a D statistic very close to zero. D-statistics larger than zero indicate a non-normal data set. A skewed data set will have a positive or negative g_1 statistic, zero indicates no skew. A positive number denotes a tail trailing off to the right, a negative number naturally denotes skew to the left. Kurtosis, the g_2 statistic, is positive if the sampled population is leptokurtic (spiked), negative if platykurtic (flat-topped), and zero if it is normal (bell shaped). The significance of the g_1 and g_2 statistics were evaluated with a simple Z test, using the variance approximations given above to calculate the standard error (in this case the standard error is the same as the standard deviation).

A few generalities may be garnered from Table 11. Exceptions to these generalizations may be found in the table; however, the overview may be more valuable than the exceptions.

Table 11A. Test for normality, skewness and kurtosis for entire Wertzville quad N = 252 points.

Dataset	Band	Mean	sd	K-S D statistic	Prob ¹ D	Skew g ₁	Prob ¹ g ₁	Kurt g ₂	Prob ¹ g ₂
Raw 1976	4	18.98	3.71	0.215	0.01	1.009	0.01	0.476	NS ²
	5	17.18	5.48	0.219	0.01	1.108	0.01	0.541	0.01
	6	57.77	8.03	0.151	0.01	-2.514	0.01	13.013	0.01
	7	62.06	11.09	0.414	0.01	-3.074	0.01	9.014	0.01
Raw 1977	4	18.74	3.49	0.183	0.01	1.068	0.01	1.229	0.01
	5	17.50	5.77	0.184	0.01	1.273	0.01	1.669	0.01
	6	55.63	7.22	0.133	0.01	-2.473	0.01	11.368	0.01
	7	58.68	11.88	0.084	0.01	-1.473	0.01	5.091	0.01
Difference	4	127.24	2.16	0.130	0.01	0.008	NS	2.031	0.01
	5	126.67	3.53	0.159	0.01	-1.301	0.01	7.514	0.01
	6	129.14	6.02	0.112	0.01	1.042	0.01	2.443	0.01
	7	130.37	8.08	0.047	0.01	1.456	0.01	4.686	0.01
Ratio	4	0.99	0.11	0.154	0.01	0.585	0.01	1.134	0.01
	5	1.03	0.19	0.161	0.01	2.111	0.01	10.153	0.01
	6	0.97	0.11	0.081	0.01	-0.137	NS	3.863	0.01
	7	0.96	0.20	0.157	0.01	4.703	0.01	50.054	0.01
1976, 7/5	1	4.00	1.42	0.118	0.01	-0.403	0.01	-0.878	0.01
1977, 7/5	1	3.79	1.55	0.117	0.01	-0.093	NS	-1.151	0.01
DR	1	4.21	0.85	0.086	0.01	1.191	0.01	2.952	0.01

¹The columns marked Prob. D, Prob. g₁, Prob g₂ list the probability that the number calculated is actually not significant. The numbers should be read with a " " sign in front of them. Hence if the number were 0.05 or smaller, then the null hypothesis (H₀: the number in question is not significantly different from zero) would be rejected at the 95% confidence level.

²NS: not significant, below 90% confidence level, i.e., the number given would be greater than 0.10.

Table 11B. Test for normality, skewness and kurtosis, forested areas only N = 146 points.

Dataset	Band	Mean	sd	K-S		Skew g1	Prob g1	Kurt g2	Prob g2
				D statistic	Prob D				
Raw 1976	4	16.40	1.24	0.210	0.01	0.881	0.01	1.908	0.01
	5	13.26	1.30	0.285	0.01	1.250	0.01	1.567	0.01
	6	58.66	5.40	0.105	0.01	0.347	0.10	1.922	0.01
	7	66.40	7.08	0.076	0.04	-0.078	NS	0.920	0.05
Raw 1977	4	16.60	1.94	0.248	0.01	2.310	0.10	9.417	0.01
	5	13.95	3.28	0.255	0.01	4.483	0.01	28.702	0.01
	6	58.14	5.48	0.078	0.03	-0.840	0.01	3.327	0.01
	7	64.56	8.31	0.110	0.01	-1.037	0.01	3.178	0.01
Difference	4	126.80	1.64	0.199	0.08	-1.078	0.01	2.124	0.01
	5	126.31	2.88	0.225	0.01	-3.995	0.01	25.549	0.01
	6	127.53	5.60	0.151	0.01	1.951	0.01	7.888	0.01
	7	128.84	8.00	0.177	0.01	2.404	0.01	10.128	0.01
Ratio	4	1.01	0.10	0.205	0.01	0.854	0.01	0.783	0.10
	5	1.05	0.19	0.215	0.01	3.097	0.01	15.670	0.01
	6	1.00	0.09	0.119	0.01	-1.283	0.01	4.705	0.01
	7	0.98	0.11	0.139	0.01	-1.720	0.01	5.993	0.01
1976, 7/5	1	5.05	0.66	0.081	0.02	-0.604	0.01	0.206	NS
1977, 7/5	1	4.83	1.08	0.137	0.01	-1.002	0.01	0.960	0.02
DR	1	4.22	0.97	0.125	0.01	1.330	0.01	2.628	0.01

1. All sample populations tested (exception: differenced data, band 4, forested), whether solely from forested areas or from the entire quad, were non-normal at the 95% level of confidence. The population characteristics responsible for this lack of normality varied.

2. Results from the forested areas were very similar to results when all cover types were considered. The Wertzville Quadrangle is heavily vegetated, most of those lands not in forest are agricultural areas that at the time of the imagery were supporting some sort of vegetation.

3. Most of the data sets, transformed or untransformed are leptokurtic. In general, Landsat data and the associated multitemporal transformations tend to have a large number of observations clustered around the band mean and a large number of outliers (relative to a normal distribution). The only data that exhibited a flat-topped distribution were the within-date band ratios (1976 and 1977 7/5 ratios).

4. Most of the data channels were skewed. The direction of skew varied; this variation in part seemed to be a function of the data transformation and whether the channel involved was in the visible or infrared portion of the spectrum.

The standard deviation threshold levels which optimized classification accuracy were surprisingly low, generally in the range of 0.50 to 1.00 standard deviations. One should remember that a specific type of change was sought; that change was spectrally manifested on only one side of the mean. The one-sided nature of the change coupled with the non-normal data produced the low threshold levels. The data non-normality may be better illustrated by comparing the actual number of flagged pixels at a particular threshold to the expected number if the population was normally distributed (see Table 12). The actual percentage of forest pixels that were noted as defoliated in the GRI was 11.68 percent.

Table 12. Comparison of actual number of changed pixels in the best single band ratio and difference images and in the Difference of Ratios image versus the theoretical percentage that would be flagged if the populations were normally distributed.

<u>Image</u>	<u>Band</u>	<u>sd Thresh</u>	<u>Side of Mean</u>	<u>Percent Forest Pixels Flagged as Change</u>	<u>Percent of Population Normal</u> ¹
Difference	5	0.35-0.60 ²	low	21.56	13.68-22.57
Ratio	5	0.65	high	19.33	24.22
Dif. of Rat.	-	1.00	high	15.05	34.13

¹A Z distribution was used to calculate this column. The normal curve areas are available in Mendenhall, 1975, Appendix 2, Table 3.

²Due to the discrete nature of the difference image, a range of standard deviation threshold levels produced the "best" classification accuracy.

Note that, in general, the percentages of alarmed pixels in the images were smaller than the corresponding values in a normal population. This

decrease mirrors the leptokurtic nature of the data. The relatively small threshold levels may have been due to the characteristic skew of the particular data sets. In each case, the data set was significantly skewed in the direction of the change being detected (see Table 11B). These results demonstrate that Landsat data collected over a diverse area such as Harrisburg is markedly non-normal, even when the cover types being considered are restricted. Certain tests and processors, such as the Bayesian classifier, assume a gaussian (or normal) distribution but are fairly robust if the data is non-normal. Analyses involving parametric statistical measures that require normality should be treated cautiously if Landsat spectral data is involved.

VI. CONCLUSIONS AND DISCUSSION

The following conclusions concerning the use of Landsat data for detecting and monitoring gypsy moth defoliation may be drawn from the literature review and the results of the study.

1. A difference in vegetative indices (such as the 7/5 ratio) classified defoliation more accurately than a strict comparison of albedo measurements, i.e., the difference or ratio transformations. The difference of ratios transformation has the added advantage of reducing an eight channel multitemporal data set to one channel. The difference and ratio transformations reduce eight channels to four.
2. Concerning differenced or ratioed data:
 - a. Band 5 is the most useful single band for discriminating canopy alteration.
 - b. Band 7 is the second most useful channel, band 4 third, and band 6 is the worst.
 - c. In all cases, additional bands did not improve change classification performance.
 - d. The ratio image produced consistently higher accuracies than the difference image. Though both compare spectral reflectance measurements obtained on two different dates, the ratio calculation produced a real number image. The discrete nature of the difference image makes it relatively insensitive to adjustments in the standard deviation threshold level for a particular band.
3. Optimal threshold levels (based on maximizing combined accuracy) were low, on the order of 0.50 to 1.00 standard deviations. One cannot associate these levels with a certain percentage of the population (via the Empirical Rule) due to the non-normality of the data.

4. All Landsat data, transformed and untransformed, were markedly non-normal. Most of the data were leptokurtic (spiked) and skewed. The direction of the skew was in part a function of the data band being considered--whether it was a visible or infrared channel--and of the transformation involved, if any.
5. The natural cubic spline function did not adequately model the relationship between the standard deviation threshold and combined accuracy. However, the natural spline predictions were much more realistic than those obtained when a quadratic model was forced to fit the empirical threshold/accuracy results. The forced quadratic regression analysis more often than not gave misleading or incorrect predictions of the optimal threshold value. The spline, a more empirical approach, produced much more reasonable modeling results, but could not handle the often, almost random behavior of the combined accuracy criterion.
6. The sequential simplex experimental design did not, in this study, adequately define the optimal threshold levels of the individual bands when used in various combinations. It did, however, produce acceptable approximations to these values. The sequential simplex design, then, should be used to define an area of the response surface that produces the highest classification accuracy for a particular band combination. The sequential simplex design provides a methodical avenue for investigating the response surface, and once a "best set" of thresholds is defined, that region of the response surface can be explored using a much smaller factorial design.
7. A combined accuracy figure, the average of the average and overall accuracies, had to be generated and analyzed. This criterion, though in itself meaningless, did combine the characteristics of average and

overall accuracy into a variable for which an optimal threshold could be defined. The combined accuracy class weights were midway between the weights associated with average and overall accuracy. Such a criterion may be useful wherever there is a need to maximize classification accuracy.

8. The moderate defoliation class was responsible for the fairly low combined classification accuracies (below 80%). Healthy forest and heavy defoliation may be correctly classified typically 85 and 95 percent of the time, respectively; moderate defoliation rarely got above 50%. Moderately defoliated areas exhibit abominable classification accuracies regardless of the data transformation used. Major increases in the classification accuracy of moderate defoliation came at the expense of healthy forest.

This study handled moderate defoliation as a valid cover type. This tended to reduce classification accuracies by alarming pixels in healthy forest. A viable alternative which would reduce this false alarm rate would be to maximize healthy forest and heavy defoliation classification accuracies to set the threshold limits. The problem with this approach is that much of the forest that lost 30-60% of its crown would go unnoticed; the amount of unhealthy forest would be underestimated.

To summarize, only three of a multitude of data transformations were tested, but the results are clear. If one is searching for a specific type of change, use a data transformation that highlights differences in the cover types being checked. In the case of forest canopy alteration, a vegetative index is most useful. Data transformations to detect change, however, are not always called for. Many have found differences in the band 5 response adequate for delineating urban boundaries. Flood or coastal zone surveys might do well to study band 7 responses at different times to detect obvious

changes in, for instance, the sizes or shapes of barrier islands or large river systems. The point is, tailor the data transformation to the specific type of change being investigated.

Second, utilize information concerning the spectral characteristics of the type of change being studied. For instance, in this project, only values to the high side of the mean of the difference of ratios data were considered as change due to gypsy moth defoliation. Low values might indicate (among other things) an increase in the biomass being sensed. Such phenomena were of no interest, since the type of change being studied concerned a loss of canopy, a reduction in the value of the 1977 7/5 ratio, and subsequently an increase in the difference of ratios value.

Finally, a review of the quantitative literature suggests that analysis of a *multitemporal* data set is inherently more accurate than classifying each date separately and comparing classifications. The errors involved with each of the separate classifications are compounded when the results are compiled into a single change image. The facts that (1) the classes represented in the statistics formulated to train the classifier may not be equivalent, and (2) misclassification errors are ubiquitous in any sort of MSS data classification, give credence to the basic tenant that post-classification differencing starts out with a marked handicap.

This research project has re-emphasized the fact that computer-aided analysis of Landsat MSS data is capable of accurately delineating healthy forest and heavily defoliated forest. Moderate defoliation causes problems; none of the change detection techniques tested accurately classified this cover type. It is unfortunate that this particular cover type class causes such problems because it is the one cover type in which the foresters and entomologists are most interested. The areas that were moderately defoliated

this year will probably be heavily infested next year. These areas, where the trees are still relatively healthy and there is a good breeding population of moths, are most deserving of the suppression efforts mentioned in the introduction. These areas cannot be reliably delineated using Landsat MSS data.

Citations

- Anderson, V.L. and R.A. McLean. 1974. "Design of Experiments, A Realistic Approach", Marcel Dekker, Inc., N.Y. 418 pgs.
- Angelici, G.L., N.A. Bryant, S.Z. Friedman. 1977. "Techniques for Land Use Change Detection Using Landsat Imagery". Proceedings, American Society of Photogrammetry, Little Rock, Arkansas. pgs. 217-228.
- Anuta, Paul. 1974. "The Use of Multispectral Difference Data for Urban Change Detection". Earth Environment and Resources Conference, Digest of Technical Papers, September 9-12, 1974, Philadelphia, PA. pgs. 116-117.
- Anuta, P. and M. Bauer. 1973. "An Analysis of Temporal Data for Crop Species Classification and Urban Change Detection". LARS Information Note 110873, Laboratory for Applications of Remote Sensing, West Lafayette, IN. 85 pgs.
- Barthmaier, E.W., T.W.D. Gregg, L.J. Sugarbaker, R.B. Scott, H.K. Hoffman, J. Roberts. 1980. Development of an Automated Method for Validating Department of Revenue Timber Harvest Reports. Pacific Northwest Regional Commission, Washington Department of Natural Resources, Olympia, Washington. 11 pp.
- Byrne, G.F. and P.F. Crapper. 1980. "Land Cover Change Detection by Principal Component Analysis of Multitemporal MSS Data: The Presence of Clouds". Proceedings, 14th International Symposium on Remote Sensing of Environment, Vol. III, San Jose, Costa Rica, April 23-30, 1980. pgs. 1375-1382.
- Chaiken, Les. 1979. "ASSESS2: Users Guide (Digital Classification Accuracy Assessment)". Unpublished manuscript, Goddard Space Flight Center, Greenbelt, MD.
- Coiner, Jerry C. 1980. "Using Landsat to Monitor Changes in Vegetation Cover Induced by Desertification Processes". Proceedings, 14th International Symposium on Remote Sensing of Environment. Vol. III, San Jose, Costa Rica. April 23-30, 1980. pgs. 1341-1351.
- Colwell, J., G. Davis, F. Thomson. 1980. Detection and Measurement of Changes in the Production and Quality of Renewable Resources. TM-145300-4-F, Environmental Research Institute of Michigan, Ann Arbor, Michigan. 144 pp.
- Davis, John C. 1973. "Statistics and Data Analysis in Geology", John Wiley and Sons, NY.
- Dottavio, C.L. 1981. Effect of Forest Canopy Closure on Incoming Solar Radiation. AgRISTARS Renewable Resources Inventory Report TM 82113, RR-61-04085, NASA/Goddard Space Flight Center, Greenbelt, MD. 24 pgs.
- Electromagnetic Systems Laboratory. 1978. "IDIMS Functional Guide", Technical Manual ESL-TM705, ESL, Inc., Sunnyvale, California, 2 volumes.
- Fisher, Sir Ronald A. 1970. "Statistical Methods for Research Workers", 14th Edition, Hafner Press, McMillan Publishing Co., NY.
- Forsythe, G.E., M.A. Malcolm, C. Moler. 1977. "Computer Methods for Mathematical Computations". Prentice-Hall, Inc., Englewood Cliffs, NJ. 259 pgs.

- Friedman, S.Z. 1978. "Progress Report for Phase 1 of the Census-Urbanized Area Application System Verification and Transfer", Earth Resources Applications Group, Image Processing Laboratory, Jet Propulsion Laboratory, 91 pgs.
- Gordon, Steven. 1980. "Utilizing Landsat Imagery to Monitor Land-Use Change, A case study in Ohio". *Remote Sensing of Environment*. 9: 189-196.
- Heller, Robert C. 1978. Case Applications of Remote Sensing for Vegetation Damage Assessment. *Photogrammetric Engineering and Remote Sensing* 44(9): 1159-1166.
- Ingram, K.J., E. Knapp, J.W. Robinson. 1981. "Change Detection Technique Development for Improved Urbanized Area Delineation". CSC/TM-81/6087, Computer Sciences Corp., Silver Springs, MD.
- Jordan, Carl F. 1969. "Derivation of Leaf-Area Index from Quality of Light on the Forest Floor". *Ecology* 50(4): 663-666.
- Joyce, A.T., J.H. Ivey and G.S. Burns. 1980. "The Use of Landsat MSS Data for Detecting Land Use Changes in Forestland". *Proceedings, 14th International Symposium on Remote Sensing of Environment, Vol. II, San Jose, Costa Rica, April 23-30, 1980*. pgs. 979-988.
- Justice, Chris. 1978. "An Examination of the Relationships Between Selected Ground Properties and Landsat MSS Data in an Area of Complex Terrain in Southern Italy". *Proceedings, American Society of Photogrammetry, Fall Meeting, Albuquerque, New Mexico*. pg. 303-328.
- Kauth, R.J. and G.S. Thomas. 1976. "The Tasseled Cap--A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat". *Proceedings, Symposium Machine Processing of Remotely Sensed Data, LARS/Purdue, West Lafayette, IN*. pgs. 4B-41 to 4B-51.
- Kauth, R.J., A.P. Pentland and G.S. Thomas. 1977. "BLOB: An Unsupervised Clustering Approach to Spatial Preprocessing of MSS Imagery". *Proceedings, Eleventh International Symposium on Remote Sensing of Environment, Vol. II, Environmental Research Institute of Michigan, Ann Arbor, Michigan*. pgs. 1309-1317.
- Kettig, R.L. and D.A. Landgrebe. 1975. "Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects". LARS Information Note 050975, Laboratory for Applications of Remote Sensing, West Lafayette, IN.
- Malila, William A. 1980. Change Vector Analysis: An Approach for Detecting Forest Changes with Landsat. *Proceedings, Machine Processing of Remotely Sensed Data Symposium, Purdue University, West Lafayette, IN*. pgs. 326-335.
- Marshall, Elist. 1981. The Summer of the Gypsy Moth. *Science*, 213 (4511): 991-993.
- Maxwell, E.R. 1976. "A Remote Rangeland Analysis System". *Journal of Range Management*. 29: 66-73.
- Nelson, Ross. 1981. "Defining the Temporal Window for Monitoring Forest Canopy Defoliation Using Landsat". 47th Annual Meeting, American Society of Photogrammetry, Washington, D.C. Feb. 22-27, 1981. pgs. 367-382.

- Pearson, R.L. and L.D. Miller. 1972. "Remote Mapping of Standing Crop Biomass for Estimation of the Productivity of the Shortgrass Prairie," Pawnee National Grasslands, Colorado. Proceedings, 8th International Symposium on Remote Sensing of Environment, Vol. II, Environmental Research Institute of Michigan, Ann Arbor, Michigan. pgs. 1355-1379.
- Riordan, Carol J. 1980. "Non Urban to Urban Land Cover Change Detection Using Landsat Data", Summary Report. Colorado Agricultural Research Experiment Station, Colorado State University, Fort Collins, Colorado, 65 pgs.
- Robinson, John W. 1979. "A Critical Review of the Change Detection and Urban Classification Literature". Computer Sciences Corporation. TM-79/6235, NASA/Goddard Space Flight Center, Greenbelt, MD.
- SAS Institute, Inc. 1979. "SAS User's Guide". Cary, NC. 494 pgs.
- Smith, Gary S. 1979. Sampling Techniques to Monitor Forest Area Change. Proceedings, 13th International Symposium on Remote Sensing of Environment, Vol. III, Ann Arbor, Michigan. pp. 1409-1418.
- Stauffer, M. and R. McKinney. 1978. "Landsat Image Differencing as an Automated Land Cover Change Detection Technique" (Interim Report). Computer Sciences Corporation Report CSC/TM-78/6215. Goddard Space Flight Center, Greenbelt, MD.
- Stauss, S., B. Wicinas and J. Ryan. 1978. "Data Stratification and the Geographic Entry System (GES): A User's Manual". Technical Memorandum ESL-TM 991, Electromagnetic Systems Laboratory, Inc., Sunnyvale, California, 224 pgs.
- Stoner, E.R., M.F. Baumgardner, and J.E. Cipra. 1972. "Determining Density of Maize Canopy: II. Airborne Multispectral Scanner Data". LARS Print 111272, Laboratory for Applications of Remote Sensing, West Lafayette, IN. 16 pp.
- Swain, Philip H. 1976. "Land Use Classification and Mapping by Machine--Assisted Analysis of Landsat Multispectral Scanner Data". LARS Information Note 111276, Laboratory for Applications of Remote Sensing, West Lafayette, IN. 414 pp.
- Todd, William J. 1977. "Urban and Regional Land Use Change Detected by Using Landsat Data". Journal Research U.S. Geological Survey, 5(5): 529-534.
- Toll, D.L., J.A. Royal and J.B. Davis. 1980. Urban Area Update Procedures Using Landsat Data. Fall Technical Meeting, American Society of Photogrammetry, Niagra Falls, New York. pgs. RS-E-1 thru 17.
- Tucker, Compton J. 1979. "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation". Remote Sensing of the Environment, 8(2): 127-150.
- Tucker, C.J. and E.L. Maxwell. 1976. "Sensor Design for Monitoring Vegetation Canopies". Photogrammetric Engineering and Remote Sensing, 42 (11): 1399-1410.
- Weismiller, R.A., S.J. Kristof, D.K. Scholz, P.E. Anuta and S.M. Momin. 1977. "Change Detection in Coastal Zone Environments". Photogrammetric Engineering and Remote Sensing. 43: 1533-1539.

- Williams, D.L. and G.F. Haver. 1976. "Forest Land Management by Satellite: Landsat-Derived Information as Input to a Forest Inventory System". Intralab Project #75-1, Goddard Space Flight Center, Greenbelt, MD. 36 pgs.
- Williams, D.L. and M.L. Stauffer. 1979. "A Forester's Look at the Applications of Image Manipulation Techniques to Multitemporal Landsat Data". Proceedings, Machine Processing of Remotely Sensed Data Symposium, LARS/Purdue, West Lafayette, IN. pgs. 368-375.
- Williams, D.L. and K.J. Ingram. 1981. "Integration of Digital Elevation Model Data and Landsat MSS Data to Quantify the Effects of Slope Orientation on the Classification of Forest Canopy Condition". Proceedings, Machine Processing of Remotely Sensed Data Symposium, LARS/Purdue, West Lafayette, IN. 11pp.

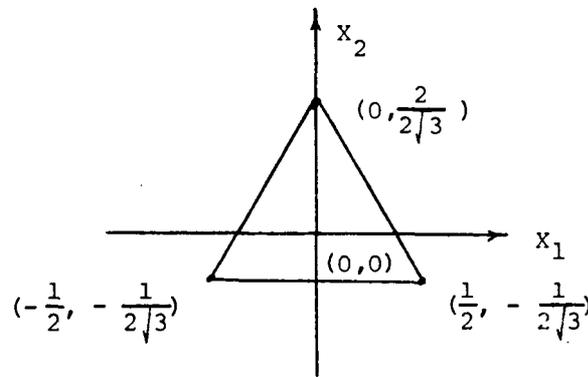
VII. APPENDIX A

(From Anderson and McLean, 1975, pg. 362-367)

Simplex (Sequential)

The sequential application of simplex designs presented here do not require that the sum of the levels of the factors add to one for each treatment combination. The number of factors, n , then, is equal to the dimension of the design. For the two factor experiment the regular simplex design is an equilateral triangle in two dimensions. It is not necessary for this design to be regular but the scaling can be done such that the same unit is used for all dimensions.

To portray the procedure we will use a two dimensional design centered around a "guessed at" maximum and call this point the centroid $(0, 0)$. The adjacent points in our design will be one unit apart. The layout for this design is as follows.



where only the three extreme points are used in the design.

The general n-dimensional design has (n + 1) points and can be arrayed as a (n + 1) by (n) matrix \mathcal{X} :

Point or treatment combination	Independent Variable (X)					
	1	2	3	4	...j...	n
1	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{6}}$	$-\frac{1}{2\sqrt{10}}$...	$-\frac{1}{\sqrt{2n(n+1)}}$
2	$\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{6}}$	$-\frac{1}{2\sqrt{10}}$...	$-\frac{1}{\sqrt{2n(n+1)}}$
3	0	$\frac{2}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{6}}$	$-\frac{1}{2\sqrt{10}}$...	$-\frac{1}{\sqrt{2n(n+1)}}$
4	0	0	$\frac{3}{2\sqrt{6}}$	$-\frac{1}{2\sqrt{10}}$...	$-\frac{1}{\sqrt{2n(n+1)}}$
.
.
.	.	:
i	X_{ij}	...
.
.
.
(n+1)	0	0	0	0	...	$\frac{n}{\sqrt{2n(n+1)}}$

Note that the two-dimensional design in the following tabulation as given previously in the upper left-hand corners of \mathcal{X} .

Point	Independent Variable	
	1	2
1	$-\frac{1}{2} = X_{11}$	$-\left(\frac{1}{2\sqrt{3}}\right) = X_{12}$
2	$\frac{1}{2} = X_{21}$	$-\left(\frac{1}{2\sqrt{3}}\right) = X_{22}$
3	$0 = X_{31}$	$\left(\frac{2}{2\sqrt{3}}\right) = X_{32}$

Next we must show how one moves sequentially to the maximum yield or response. Let us assume that the point $(0, 2/(2\sqrt{3}))$ gives the minimum yield of the three points in the two-dimensional design. The sequential method is to delete that point and form another simplex with the remaining two points and a point on the opposite side. In this case the new point would be $(0, -4/(2\sqrt{3}))$. The algebraic method for finding the new point is to delete the minimum point from the array giving the remaining array as follows.

Point	Independent Variable	
	1	2
1	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$
2	$\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$

The next step is to add the levels of each factor and multiply each by $2/n$ (where $n =$ the number of X 's) or $2/2 = 1$ for this case, and subtract the level of the factor deleted. For variable 1 we have

$$\frac{2}{2} \left(-\frac{1}{2} + \frac{1}{2} \right) = 0$$

and subtract 0 giving 0. For variable 2

$$\frac{2}{2} \left(-\frac{1}{2\sqrt{3}} - \frac{1}{2\sqrt{3}} \right) = -\frac{2}{2\sqrt{3}}$$

and subtract $2/(2\sqrt{3})$ which gives $-4/(2\sqrt{3})$.

Using the X_{ij} notation of the $(n+1)$ by n matrix given previously to denote the level of the j^{th} variable of the i^{th} treatment combination, we can find the new level X_{ij}^* by using

$$X_{ij}^* = \left(\frac{2}{n} (X_{1j} + X_{2j} + \dots + X_{i-1,j} + X_{i+1,j} + \dots + X_{n+1,j}) - X_{ij} \right)$$

The procedure to find the optimum region of the treatment combinations is to continue to delete the point or treatment combination that has the lowest

response in the simplex, replacing this point by a point opposite to it as described above. This procedure is followed until an optimum region is located. The optimum region is indicated by several new points, say n , failing to produce a response that is greater than that obtained at a previous treatment combination. At this point the optimum treatment combination should be repeated before experimentation is stopped in order to assure that this point is truly optimum.

If in the process of stepping through the factor space with successive simplexes, the yield of a new point is less than any other point in the current simplex, do not go back to the point that was previously vacated. In this case determine the next lowest point and move opposite it. This makes up a different simplex which is the basis for the next decision.

This concludes the design discussion. The analysis, in addition to finding the optimum point, may include running a regression analysis on those points near the optimum.

An example of the procedure is the following electrical engineering problem in which the response y is output power and the inputs are (X_1) voltage, (X_2) relative humidity, and (X_3) temperature. The present system for X_1 is 110 V, X_2 is 43%, X_3 is 86°F with $y = 58\%$. Hence the point

$$(0, 0, 0) = (110, 43, 86)$$

has a yield of 58. For X_1 , X_2 , and X_3 the experimenter felt that increments of 10, 2, and 0.5, respectively, were equally important. Using these increments to code about the centroid, we obtain the four treatment combinations

1.	105,	42.4,	85.8
2.	115,	42.4,	85.8
3.	110,	44.2,	85.8
4.	110,	43.0,	86.6

from the following matrix:

$$\begin{bmatrix} -\frac{1}{2}, & -\frac{1}{2\sqrt{3}}, & -\frac{1}{2\sqrt{6}} \\ \frac{1}{2}, & -\frac{1}{2\sqrt{3}}, & -\frac{1}{2\sqrt{6}} \\ 0, & \frac{2}{2\sqrt{3}}, & -\frac{1}{2\sqrt{6}} \\ 0, & 0, & \frac{3}{2\sqrt{6}} \end{bmatrix}$$

The yields for the four treatment combinations were 52, 62, 61, 57, respectively. Hence, delete treatment combination 1 and replace it by the three levels:

$$\text{Level 1} = \frac{2}{3} (115 + 110 + 110) - 105 = 118.3$$

$$\text{Level 2} = \frac{2}{3} (42.4 + 44.2 + 43.0) - 42.4 = 44.0$$

$$\text{Level 3} = \frac{2}{3} (85.8 + 85.8 + 86.6) - 85.8 = 86.3$$

This new treatment combination (118.3, 44.0, 86.3), along with the remaining three treatment combinations, forms the new simplex. After running the new treatment combination, the sequential procedure continues as discussed previously.

The efficiency of sequential designs is measured by the number of points required to reach the maximum. For a given surface, one can determine the rate of advance to the maximum. Spendley, Hext, and Himsworth (1962) show that as the standard deviation (variation of the response for the same treatment combinations) increases the rate of advance decreases as $(1/(\text{standard deviation}))$. For example, the rate of advance in an experiment with standard deviation of 3 is twice that of an experiment with standard deviation of 6.

On comparing achievement of sequential procedures to reach maximum, Spendley, Hext, and Himsworth (1962) indicate that the simplex design is

second only to the steepest ascent (regression) procedure if there is no error and somewhat poorer (but not too bad) when error is present.

Box and Behnken (1960) describe a "simplex-sum" design, a type of second order design, that combines the vectors defining the points of the initial simplex in pairs, threes and so on. One can use the midpoints, or points off the midpoints, of the vectors joining the vertices to estimate the curvature.

Box, G.E.P. and Behnken, D.W. Ann. Math. Stat. 31: 838 (1960).

Spendley, W., Hext, G.R., and Himsworth, F.R. Technometrics 4: 441 (1962).

Appendix B - Contingency Tables
 Best Methods for Delineating Defoliation;
 Difference, Ratio, and Difference of Ratios Images

Contingency tables provide the data necessary to formulate various classification accuracy criteria. Descriptive statistics such as the average, overall and combined accuracies presented in the text may be calculated from these tables. Readers reviewing this work may be interested in more specific or entirely different accuracy indicators. Hence the best approaches for the difference, ratio and difference of ratios images are given below. These images correspond to those listed in Table 10 of the text.

The Landsat classification noted a forest pixel as either changed or unchanged. Likewise the accuracy figures were calculated as if the ground reference image (GRI) was composed of only those two classes. Change - no change on the Landsat images correspond to healthy forest/defoliated forest on the GRI. The defoliated forest category was composed of moderately and heavily defoliated areas. These labels are maintained in the contingency tables so that the reader may note which ground reference class is most error prone.

Difference Image, band 5. Standard deviation threshold 0.35-0.60

		Ground Reference Image			
		No Change Healthy	Change		Mod and Hvy
			Moderate	Heavy	
Landsat	No Change	25910 ¹ 83.4 ²	1657 50.1	23 2.9	1680 40.9
Class.	Change	5157 16.6	1650 49.9	778 97.1	2428 59.1
Total		31067 100.0	3307 100.0	801 100.0	4108 100.0

1. Number of pixels

2. Percent of total pixels in category

Example: Classification accuracy calculations (as noted in Table 4b, text)

Healthy: $(25910/31067) \times 100 = 83.4\%$
 Moderate Def: $(1650/3307) \times 100 = 49.9\%$
 Heavy Def: $(778/801) \times 100 = 97.1\%$
 Change: $((1650 + 778)/4108) \times 100 = 59.1\%$
 Average Accuracy = $(\% \text{ Correct No Change} + \% \text{ Correct Change})/2$
 = $(83.4 + 59.1)/2 = 71.3\%$
 Overall Accuracy = $(\# \text{ correctly classified pixels}/\# \text{ test pixels}) \times 100$
 = $((25910 + 1650 + 778)/(31067 + 3307 + 801)) \times 100 = 80.6\%$
 Combined Accuracy = $(\text{Avg Acc} + \text{Over Acc})/2$
 = $(71.3 + 80.6)/2 = 75.9\%$

Ratio Image, band 5, standard deviation threshold 0.65

		No Change Healthy	Ground Reference Image Change		Mod and Hvy
			Moderate	Heavy	
Landsat	No	26675	1674	27	1701
	Change	85.9	50.6	3.4	41.4
Class.	Change	4392	1633	774	2407
		14.1	49.4	96.6	58.6
Total		31067	3307	801	4108
		100.0	100.0	100.0	100.0

Average Accuracy: 72.2%
 Overall Accuracy: 82.7%
 Combined Accuracy: 77.5%

Difference of Ratios, standard deviation threshold 1.00

		No Change Healthy	Ground Reference Image Change		Mod and Hvy
			Moderate	Heavy	
Landsat	No	27894	1963	26	1989
	Change	89.8	59.4	3.2	48.4
Class.	Change	3173	1344	775	2119
		10.2	40.6	96.8	51.6
Total		31067	3307	801	4108
		100.0	100.0	100.0	100.0

Average Accuracy: 70.7%
 Overall Accuracy: 85.3%
 Combined Accuracy: 78.0%

BIBLIOGRAPHIC DATA SHEET

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Detecting Forest Canopy Change Using Landsat		5. Report Date April, 1982	
		6. Performing Organization Code	
7. Author(s) Ross F. Nelson		8. Performing Organization Report No.	
9. Performing Organization Name and Address Earth Resources Branch, Code 923, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771		10. Work Unit No.	
		11. Contract or Grant No.	
		13. Type of Report and Period Covered Technical Memorandum 10/1/80 - 9/30/81	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract Multitemporal Landsat multispectral scanner data were analyzed to test various computer-aided analysis techniques for detecting significant forest canopy alteration. Three data transformations, differencing, ratioing, and a difference of ratios, were tested to determine which best delineated gypsy moth defoliation. Response surface analyses were conducted to determine optimal threshold levels for the individual bands and band combinations. Results indicate that, of the three transformations investigated, a difference of ratios (band 7/band 5) transformation most accurately delineates forest change due to gypsy moth activity. Band 5 (0.6-0.7 um) ratioed data did nearly as well, however, other single bands and band combinations did not improve upon the band 5 ratio results.			
17. Key Words (Selected by Author(s)) Change detection Forest canopy Insect Defoliation Landsat MSS		18. Distribution Statement	
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*